THE UNIVERSITY OF ALBERTA

THE INTERACTION OF TONE AND INTONATION IN MANDARIN CHINESE

by

(C) BRUCE A. CONNELL

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

IN

SPEECH PRODUCTION AND PERCEPTION

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA

FALL, 1982

Abstract

Until recently, the linguistic phenomemon of tone has not enjoyed great popularity as an area of study. This is especially true from the point of view of experimental phonetics, and specifically with regard to the interaction between tone and other prosodic phenomena.

Some research has been done on the productive aspects of this interaction, and evidence indicates that tones can be perturbed or manipulated as a result of the effects of intonation. Very little work has been done from the perceptual standpoint and, given that tones can be changed both in shape and register, the important question arises as to what effect this perturbation has on the perception of tones; or, how far can the shape of a tone be changed before it is consistently recognized as a different tone?

This thesis reports on experimentation designed to provide at least a tentative answer to the above question. Naturally produced syllables were artificially adjusted by computer to simulate the effects of intonation on sentence final syllables of Mandarin Chinese. These were then played back for a group of native Chinese speakers who judged them for recognizability and classified them accordingly.

Results indicate that perceptual confusion can arise and that these confusions correspond generally with expectations based primarily on production data. There is apparently a wide range of leeway, however, before these confusions do occur.

Acknowledgements

I would like to acknowledge, and express my gratitude to, those who assisted, in a large or small way in the completion of this thesis. First, and foremost, to my supervisor Dr. J. T. Hogan, who saw this through from the beginning and provided invaluable assistance, both directly and indirectly related to the work at hand. I am also greatly indebted to the other members of my committee, all of whom provided critical suggestions at appropriate moments. They are Dr. A. J. Rozsypal and Dr. W. J. Baker both of the Department of Experimental Linguistics, and Dr. Stan Munro, of the Department of South-East Asian Studies. A special note of thanks must also go to two other members of the Department of Experimental Linguistics; Allan Opperthauser, departmental technician and wizard who helped immeasuably with technical matters of which I knew nothing, and my colleague Justin Tse Chen, who gave essential assistance on matters pertaining to the Chinese language, and who organized the subject pool. Thanks, of course, must also be extended to all of the people who partipated in the experiment, exchange scholars from the People's Republic of China, who co-operated smilingly and without the usual financial enticement. I am indebted also to all those friends and colleagues in the department who helped out in so many other ways whether, they realized it or not. Finally, and in some ways most importantly, to the Department of Experimental Linguistics for the financial

support that made the entire venture possible.

Table of Contents

Introduction

This thesis is a presentation of an experimental investigation into the perturbability of the tones of Modern Standard Chinese (MSC), or Mandarin. Since a change in tone in MSC will change the meaning of a word, it is seen as interesting and important to discover how much the shape of an individual tone can vary before it is recognized as a different tone.

Very little research has been done on this topic. A review of the relevant experimental work is presented in Chapter One, following an introductory section dealing with definitional and background concerns.

Chapter Two presents details on the methodology used in constructing the stimuli and running the experiment. Also included is information pertaining to the subject pool, acoustical data concerning the stimuli, and a sample page from the answer sheet.

Chapter Three is a statistical analysis of the results obtained. A brief description and justificaton of the statistical technique used (hierarchical clustering, or cluster analysis) is given, as is a presentation of what this procedure revealed. Graphs are included.

Chapter Four is a discussion of the results, as revealed by the cluster analysis; much of this discussion focusses on differences among the subjects; consequently, attention is paid to background information about the subjects. Identification curves depicting the responses of a

particular sub-set of the subject pool are also presented and discussed in this chapter. This sub-set consisted of ten subjects who were mostly native speakers of Mandarin. The group was apparent through studying the raw data, and the fact of its existance was supported by the cluster analysis.

Chapter Five is an evaluation of the results in terms of other research done as presented in Chapter One. Also, an evaluation of the present research is offered, with suggestions for improvement, and directions for future research are included.

Chapter Six provides a brief summary of the thesis, and is followed by an appendix containing the computer programs used in the course of the research, a map of China indicating the major linguistic groupings and, finally, a bibliography of sources consulted.

I. LITERATURE REVIEW

The phenomena of tone have, in the past decade, enjoyed a considerable upsurge in popularity as objects of study. This may be due to recognition on the part of linguists of the fact that the majority of the people in the world speak a tone language. (Tone languages are found in virtually all geographical regions of Earth, but primarily in south-east Asia and in Africa.) Or perhaps there is a new breed of linguist, more adventurous and looking for uncharted territory to investigate. Current interest in tonal phenomena runs the entire range of linguistic sub-sets, i.e., from phonetic considerations, to phonological, and to syntactic and typological interests. Phonetic aspects of research generally consist of measurement and perceptual experiments; phonological work consists primarily of rule writing to describe the occurrence and co-occurrence of tones, i.e., how they interact in the presence of each other, usually referred as tone sandhi; and in some languages tone has a syntactic function, which is also being examined.

In connection with typological research being done, there is currently controversy concerning what should properly be considered a tone language. The broadest classification would include any and all languages using tone in a linguistic fashion (lexical or grammatical) as opposed to extra- or paralinguistic uses (expressive or attitudinal). This would allow the inclusion of languages

such as Serbo-Croatian and Swedish, as well as other more easily classified languages, such as the Chinese and Bantu languages. A more restricted definition of tone language includes only those languages in which tone functions lexically, for example, all languages and dialects grouped under the name Chinese. A detailed examination of the typological problems of tone languages is of interest, but beyond the scope of this thesis; for further discussion, the reader is referred especially to Pike (1948), Lehiste (1970), and McCauley (1978).

At this point, it may be best to describe what is meant by tone, as differentiated from other prosodic phenomena, especially intonation. Tone generally refers to the contrastive, or linguistic, functioning of the fundamental frequency at the word or syllable level; intonation is the contrastive functioning of the fundamental frequency at the sentential level. It has also been argued (Jacobson, Fant and Halle, 1952)that intonation may fulfill an organizational function for the sentence. Physiologically speaking, then, the correlate of both tone and intonation is the frequency of vibration of the vocal folds during phonation. This is not to suggest that pitch is the only acoustic cue for tone and intonation; both duration and intensity play a part as well, and the relative importance of the three will vary with the situation. Pitch, however, is generally accepted as having the greatest importance, with duration and intensity usually serving as auxiliary

information.

In opposition to tone languages there are what are often referred to as intonation languages. This, however, is a misnomer since virtually all languages of the world use intonation linguistically. Bolinger (1978), in his survey, mentions only one language, Amahuaca (eastern Peru), that has been described as having no linguistically contrastive intonations. Consequently, the preference here is for reference to tone languages and non-tonal languages.

Intonation, as with tone, has been little studied until recently. Many of the reasons for this also apply to the study of tone. The problems involved with studying intonation are seen in the difficulty with determining exactly what constitutes intonation; how to distinguish the linguistic uses of intonation from the paralinguistic, in determining what is the relative importance of intonation vis-a-vis other linguistic devices (e.g., the syntactic use of particles, see below, p. 12), and in establishing the relative importance of different types of intonation.

While most linguists may have an intuitive notion of what intonation is, a glance through the literature turns up widely varying definitons. These definitons range, from the simple (as above) "linguistically relevant use of pitch at the sentential (but actually clausal) level", to the more complex, such as that of Crystal (1969), who defines intonation as:

> not a single system of contours, levels, etc, but as a complex of features from different systems...tone, pitch-range, and loudness, with rhythmicality and tempo closely related (p. 195).

While preferring an analysis closer to that of Crystal than the less complex one offered above, for the purposes of this study, the simpler definition (i.e., the linguistically relevant use of pitch at the sentential level)is sufficient. This is possible for a number of reasons, chiefly in that pitch, or changes in fundamental frequency, as discussed above, are generally accepted as being the main manifestation of tone or intonation; this is especially true of the particular environment examined in the experiment contained in this thesis. The difficulty surrounding the differentiation of linguistic and paralinguistic uses of intonation (or tone) is also not pertinent for this study, since in either usage, intonation, when manifested as a pitch phenomenon, is the result of the same physiological activity (i.e., the frequency of the vibrating of the vocal folds).

It is this physiological activity that is the root of the present problem for, as mentioned, tone also is the result of the use of change in fundamental frequency. Immediately one can see the potential for problems when thinking of intonation in context of tone languages, regardless of the type of tone language. A number of questions arise; which takes precedence, tone or intonation;

how is ambiguity resolved, if it occurs; does the relative importance of different acoustic cues for tone vary in context of different intonations; and so on. Essentially, these questions deal with the interaction between tone and intonation, and obviously there must be an interaction; and they can be investigated from a variety of points of view, from phonetics to semantics.

It is generally observed that intonation will take precedence over tone; that is, tones will be modified by intonation, rather than the reverse (Chao, 1968). One example of this is downdrift, a phenomenon present in many African tone languages. In this situation, a high tone near the end of a declarative phrase will be realized not only at a lower pitch than it was at the beginning of the sentence, but also possibly lower than a low tone appearing early in the sentence.

While there has been an increasing amount of research done on tone languages, including at the level of experimental phonetics, this work has focussed on tones in 'citation' form, i.e., as spoken in isolation, or on tone sandhi. ('Sandhi' is a term drawn from morphology which refers to joining processes; in this context, 'tone sandhi' refers to the influence adjacent tones have on each other.) Very little research has actually been done on the interaction between tone and other prosodic phenomena, most important of which is intonation.

Before describing in detail the research done on this for Chinese, and that relevant to the present research, a brief discussion of certain of the more general observations will be presented. Most important of these observations is that intonation will not usually change the shape of a tone beyond recognition. (Chinese is what Pike (1948) calls a contour tone language, where pitch movement and direction are critical, as opposed to register tone languages, where tones are generally level in citation form, but the relative height is of importance.) What has most generally been observed is that, rather than the shape of the tone being drastically altered, the entire tone may be raised or lowered. This, however, is not the case in all contexts. An environment where the shape of the tone is observed to change is sentence-final position, where distinctive intonations may be manifested later in the syllable, after the tone (Chao, 1968), or may alter the shape of the tone. This alteration may, for example, see a rising tone lowered by a falling intonation, to the point of approximating a level tone. It has also been generally noted that sentence final intonations will be manifested primarily by pitch, whereas in other sentence environments there may be greater use of duration or intensity.

One of the earliest instrumental investigations of tone and intonation in Chinese was Chang (1958), done on the dialect of Chengdu, spoken in the province of Szechuan. In Chengdu a sub-dialect of Mandarin is spoken; however, it

varies from Modern Standard Chinese, (MSC, the dialect of the Peiking area) in its tonal system. Both have four tones, however Chengdu tones are: 1) mid-rising, 2) mid-falling, 3) high-falling, 4) low falling-rising. Opposed to these are MSC 1) level, 2) rising, 3) falling-rising, and 4) falling.

The importance of Chang's work is perhaps diminished by the fact that it is based on data from one aged informant, and by primitive research techniques; it is however, of interest in that her results have been corroborated by other, more sophisticated research, in at least a general fashion. (See below.)

Regarding intonation, Chang determined three important factors: 1) the pitch level on which the sentence is spoken; this she divides into five levels; 2) the pitch range, divided into wide, medium and narrow; and 3) the effects on the final syllable. The pitch level is claimed to have a definite relationship to the type of sentence. Chang has distinguished seven sentence types based on intonation, and expressing different emotions or attitudes, such as annoyance, surprise, or contempt. It is also claimed that pitch range can therefore be a clue to the emotional state of the speaker. The narrow range was observed to have the tendency to flatten rising or falling tones.

The perturbation of tones in sentence final position occurs according to whether the intonation is of a rising or falling nature. These are summarized in Chang, p. 83, but briefly, rising tones are made level under the influence of

a falling intonation, and falling tones are levelled by a rising intonation. The perception of rising and falling tones apparently remains unaffected by concommitant changes in intonation.

These two 'tunes' are used for different sentence types and together with changes in level and pitch range are seen to constitute the intonation of the sentence. Chang is not explicit on this, though she does state that the type of intonation is indicated by the final syllable. The type of intonation (i.e., rising or falling) associated with different sentences types by Chang does not entirely match the descriptions offered by Ho (1977).

Vance (1976) suggested that changes in sentence final contours might not be the result of a tone/intonation interaction, but might perhaps be conventionalized. Kratchovil's (1968) findings on Mandarin regarding the effects of stress on tones, according to Vance, suggest that each tone has a stressed variant. However, Vance argues that this would not be the case for sentence finals in Cantonese, since there is no tone sandhi in Cantonese in this environment. So the possibility of intonation interacting with tone does exist, and would be especially problematic in Hong Kong Cantonese, since this dialect has four level tones. Vance, therefore, designed experiments to: 1) test for possible effects of intonation on tones; and 2) see if predictions concerning possible confusions, based on the first experiment, would actually be borne out in a

perceptual experiment.

The first experiment involved three different sentence contexts: 1) with the test word in medial position of a declarative sentence; 2) placing the test word in final position of a declarative sentence; this context involved two sentences, the first of which was to ensure that the test word was old, or given information, thereby minimizing the possibility of contrastive stress being used. The third context placed the test words in a contrastive situation.

This experiment did show that intonation can have a lowering effect on sentence final tones in Cantonese; this was the case for four of the five subjects. The fifth subject apparently had some difficulty in comprehending the experiment. Contrastive stress, however, seemed to have only a minor effect on the fundamental frequency, and in all cases relative pitch relationships were maintained. Based on the assumption that relative pitch is the sole, or at least primary, cue for tone in Cantonese, Vance proceeded with his second experiment, to test for confusability of tones. The stimuli for this experiment were made by tape-splicing, to interchange test words from the different contexts in experiment one.

In general, the results bore out Vance's predictions about confusability, and relative pitch did seem to be the primary cue for tonal distinctions. Not all of his results are so easily explained, however. In the case of low tones there could well be other cues interacting,such as duration,

or that subtle differences in the fundamental frequency contour might have allowed for easier discrimination. This latter notion is first suggested by Abramson (1962) as a possible solution for a problematic aspect of tonal phenomena in Thai, however his work was done on citation forms. As Vance points out, these subtle variations could well disappear in connected speech. It is likely that any strong conclusions should be postponed pending further research, since the number of tokens used in Vance's experiment are too few to use as a basis for conclusive judgements.

Vance's results do indicate that sentence final tone lowering does occur in ordinary declarative sentences, and he suggests that Lieberman's (1967) breath-group hypothesis would account for this. (The breath group hypothesis suggests that sentence or clause final lowering of the fundamental frequency is a result of a decrease in the subglottal pressure. It is therefore considered that this phenomenon is the result of innate physiological conditions. For elaboration, see Lieberman (1967); for a critique of this hypothesis, the reader is referred to Ohala (1978).) His results also indicate that this lowering may have an effect on intelligibility, though tonal distinctions are probably not usually neutralized. It is obvious from Vance's research that more work needs to be done, particularily on the perceptual aspect of this phenomenon.

Research on Mandarin

One of the more prolific researchers in this area has been Aichen Ting Ho (1976a, 1976b, 1977). Ho has carried out work on citation form tones, tones interacting with different types of intonation, and the effects of grammatical structure on the realization of tones.

Ho (1976a) contrasted the four tones of MSC in different segmental and grammatical environments. The effects of tone sandhi were controlled by placing a fourth tone word (usually /rien/) before or after the test words, depending on sentence structure. A corpus of six sentences was recorded by five Mandarin speakers. Measurements were done for fundamental frequency at three pointsof the syllable nucleus, beginning, middle, and end, (and any inflection points), as well as for duration. Means were calculated and compared for each of these environments

Overall, it appeared that both intonation and grammatical function have a definite influence on tonal contours but, as before, the basic shape of the contours, i.e., level, rising, falling-rising, and falling, are more or less maintained. In environments with a falling intonation (sentence or clause), tone contours of test items were found to have a higher fundamental frequency sentence initially, somewhat lower medially (clause final), and lower still for sentence final syllables. This again would fit in with Lieberman's breath-group theory.

The influence of rising and falling intonations in sentence final position was also examined in Ho (1976a). Strictly speaking, the examples of rising intonation were not sentence final, as phrases having a final question particle were used. Therefore, the comparison is between X___# (falling, declarative) and X___Q# (rising, question). Analysis of the fundamental frequency showed that all tones except Tone 3 started and ended higher in frequency in the environment of a rising intonation. Tone 3 also started higher, but ended lower in this environment, and the low point of the dip was considerably higher.

The general findings here are not in accord with Chao (1968, p.44) who claims that these intonations affect only the last syllable, and even in that syllable appear to be added after the tone. (In fact, Chao considers these two sentence final contours to be grammatical particles.) Nor do they agree with Abramson's (1979) findings for Thai, which indicate that the "terminal juncture" is carried primarily by the particle, where such is present.

In the environments where Ho considered grammatical functions, she found the fundamental frequency to be higher for all tones in an environment not containing an emphatic modifier as compared to sentences with an EMPH (X___#vs. EMPH X___#). One other instance of grammatical influence was considered, that being the effect of a preceding demonstrative. The two environments compared here were # DEM___X vs. #___X. In this case the change in fundamental

frequency was not uniform for all tones. Fundamental frequency contours for Tones 1 and 4 start higher when preceded by a demonstrative than when not; however the reverse is true for Tone 2 and Tone 3 i.e., the fundamental frequency starts lower when preceded by a demonstrative. The range of the fundamental frequency is therefore wider for elements following a demonstrative than those in sentence initial position.

Ho also found syllable duration to be affected by grammatical structure. For example, in all cases duration was greater for elements preceded by a grammatical marker (whether EMPH or DEM) than those not. In the case of falling and rising intonations in sentence final position, Tone 1 and Tone 4 were longer when associated with a falling intonation, whereas Tone 2 and Tone 3 were longer in the rising environment. Perceptual experiments were not done to determine whether changes in fundamental frequency or in duration were the more relevant cues; it is assumed on the basis of other research (as discussed above) that pitch is the more important cue.

Ho (1976b) deals primarily with citation forms, attempting to determine the amount of tonal variation associated with the segmental structure of the syllable. Also examined, however, is the effect of sentence environment, which was compared to that of segmental considerations. The important results reported here deal first with tone range, then with the effect of the vowel and

preceding consonant. ('Tone range' refers to the entire fundamental frequency range in which a given tone may be produced.) Words in citation form were found to have the widest tone range, in the order of Tone 1 to Tone 4, with Tone 4 having the widest range. This order changed somewhat in sentential environments, however it was the location of the tone on the frequency scale that varied more, rather than the actual range of the tone.

Regarding the degree of influence of the three factors, preceding consonant, vowel, and sentence environment, on tonal shape, it was determined that sentence environment has the greatest effect, followed by vowel type and then preceding consonant.

Ho (1977) reports on investigations into the acoustic parameters of three types of intonation: declarative, interrogative, and exclamatory, and examines the influence of intonation on the four tones in sentence-final position.

To do this, a sentence frame of five words, "tse ke tsi si nian ___", (This word is pronounced ___"), was used; this choice was governed by the fact that the sentence can be said with each of the three intonations, and also that each word carries a tone four (except 'ke', which is atonic), thereby neutralizing the possibility of interference from tone sandhi. To investigate the effect on the final syllable, two CV-structure syllables were used, /ki, pa/ . These were chosen on the basis that both exist as distinct lexical items using each of the four tones. The sentences

(24 in all) were recorded by six subjects. The test items (all words in all sentences, as well as the sentence final words in citation form) were then measured for fundamental frequency at three points: beginning, middle, and end; for duration, and for amplitude. Mean values were then calculated and compared.

Ho's data again shows the fundamental frequency of tones to be influenced by intonation. The exclamatory intonation had the effect of raising the fundamental frequency for the entire sentence, as did the interrogative intonation, though to a lesser extent. The overall fundamental frequency was lowest for declarative sentences. Exclamatory and interrogative intonations also had the effect of raising the fundamental frequency in sentence-finals, whereas the declarative lowers the fundamental. The fundamental frequency was also found to be modified by word position, as were tone range, amplitude, and duration. These latter were also affected by intonation.

Concerning duration, it was found that position within the sentence had a greater effect than did sentence type (i.e., intonation). However, durational differences between the three sentence types were greatest for words in final position.

Much of Ho's data consists of average values of the four tones. Since the tones are all different and distinct, these values are meaningless. However, she also presents an analysis of the individual tones, as affected by the three

intonations in sentence final position, and compares these to the words in citation form. From this viewpoint, it is seen that intonation has definitely modified the tones, both in relative height and in shape. For all tones, the relative height, by type of intonation, is as follows: exclamatory is higher than interrogative, though only slightly; this in turn is higher than the citation form, which is higher than the declarative. This is the case throughout, except for tone four, where the endpoint of the exclamatory is below that of the interrogative. The difference in relative height is greater in all cases at the end than at the beginning, showing that intonation has also had an effect on the shape of the tones, though Ho points out that the contours do not change greatly from the citation forms. In this experiment then, some small measure of support is found for Chao's (1968) assertion, detailed above (p.6).

One final, and potentially serious, criticism of Ho's experiment pertains to her sentence frame and test words. While she has claimed to be working on Mandarin, her carrier sentence appears to be Taiwanese; since she has not stated which orthographical system she has used, this is hard to verify. The problem is compounded by her two test words, /ki/ and /pa/, of which /ki/ certainly is not Mandarin.

A relatively large amount of research has been conducted in the Soviet Union, most of which is unavailable in the West. Rumjancev (1972) conducted a number of experiments on tone, intonation, and the interaction of the

two. Lyovin (1978) subjects this book, which has never been translated or made available in the West, to a lengthy review for the benefit of those among us who, for one reason or another, do not have access to the Russian original.

The first part of Rumjancev's book deals with tones in their citation form, and will be discussed here only briefly; the second part reports on a number of experiments investgating the interplay of various prosodic features, the characteristics of different intonations, as well as the interaction of intonation and morphological devices for marking syntactic structures.

By varying the duration of citation syllables, the following confusions of tone identity were evident in perceptualexperiments: Tone 4 is confused with Tone 1; Rumjancev suggests this is due to both of them being in the same register, at least for their starting point. Tone 4 is more easily recognized than Tone 2, because it falls more rapidly than Tone 2 rises. Tone 2 can be confused with Tone 1, especially in case of shorter durations, since the rise is then minimal. Tone 3 is sometimes confused with Tone 4, since the initial portion of Tone 3 is similar in shape to Tone 4, even though of a different register; amplitude apparently played a critical role in this case. Tone 3 is also at times mistaken for Tone 1; this occurs in cases where Tone 3 was lacking its initial, falling portion.

By means of recognition experiments based on the erasure of initial consonants it was concluded that these,

although voiceless, could also possibly contain cues for tone identity. Rumjancev unfortunately does not provide detail on the nature of the stimuli used; Lyovin, however, speculates that the transitional portion of the syllable, between the initial consonants and the vowel, may contain some kind of cue despite its briefness. Support for this notion is offered, in that cues for place of articulation of the preceding consonant can be found in this transition. A general conclusion at this point is that cues to tonal identity may be contained in all parts of the syllable. Register features, however, should be considered unreliable out of context, since these are relative to the individual voice. ('Register' refers to the overall pitch level of the sentence.)

Rumjancev disputes the notion that intonation is solely, or even primarily tied to pitch, and contends that amplitude and duration can both play a prominent role, particularily in tone languages, where pitch is the main cue for the tones. (It can be added here that the existence of particles in tone languages such as Chinese and Thai, to express grammatical functions and attitudinal concerns, also helps to lessen the load on 'pitch-oriented' intonation).

In terms of sentence intonation in MSC however, Lyovin reports Rumjancev's general observations as being quite close to those noted above. Specifically, when the pitch features of intonation mix with those of lexical tone, the contours of tones are basically left intact; register

features are, however, significantly affected. Also noted is the occurrence of "post-tonal stretches of phonation," which corresponds with Chao's observation that intonation may occur after the tone of the syllable (see Chao, 1968, p.812).

Although Rumjancev has apparently argued that duration and amplitude are important cues for intonation, in terms of sentence final intonation, and specifically interrogative and declarative intonations, he does confirm that pitch is the primary cue (Lyovin, 1978, p.148). Here again it is argued that basic contours are not normally significantly affected, but register and steepness of rise or fall of contours are. Furthermore, the significant part of the intonation is localized in the last accented syllable (i.e., tonic syllable) and reduced syllable, (i.e., particle, when present). This was determined by interchanging the final word of declarative and interrogative sentences. Sentences were judged by native listeners as being interrogative or declarative, strictly on the basis of the last word. Of importance also is that these manipulated sentences were judged to be natural by the subjects.

Rumjancev also looked at what Lyovin has characterized as "nonterminal" versus "utterance terminal" intonation, and the effects of these on tones. Non-terminal intonation often is the marker of subordinate, dependant clauses; utterance terminal applies to independent clauses. In MSC dependent clauses are obligatorily marked as such morphologically, and

can therefore also stand as independent clauses. This facilitated experimentation.

A number of types of dependent clauses were examined and compared to independent clauses. Among the dependent clauses, very little difference was found in intonation; all had the same "nonterminal" intonation features. These features occur on the final word and are: increased duration, increased amplitude, and raised pitch. It was also determined that the presence of a conjunction with the dependent clause can weaken, but not neutralize, the nonterminal intonation, i.e., the intonation of the subordinate is never identical to that of the principle clause. Furthermore, in terms of the acoustic characteristics, especially pitch, very little difference was found between this non-terminal intonation and the interrogative intonation described above.

As for the effects on pitch features of lexical tones, the utterance terminal intonation apparently causes the greater degree of distortion. The pitch lowering effect of this intonation, "very often neutralizes the contrast between Tone 2 and Tone 3 syllables."(Lyovin, p.153).

A number of other interesting questions are treated by Rumjancev, for example whether or not the question particle 'ma' would override or weaken interrogative intonation. (Chinese makes use of a number of different particles to fulfill syntactic functions; in this particular instance, 'ma' acts in much the same manner as the sentence final 'eh'

in Canadian English.) Different experimental techniques were used to try to arrive at an answer to this question, and opposing solutions resulted. Lyovin points to methodological difficulties in all cases, and although Rumjancev argues for the primacy of intonational markers, it is apparent that a more sophisticated experimental technique is needed.

Lyovin points out a number of other faults with Rumjancev's book. The most serious of these are the ommission of details on methodology for most of the experiments reported, and the absence of information pertaining to the subjects used. Both types of information are essential for anyone interested in replicating Rumjancev's work.

Rumjancev's findings are important and of interest, though, in that they do confirm or correspond to the results of others; Rumjancev has also broken considerable new ground in his research, and left direction for much work to be done.

Whether pitch is always the main characteristic of intonation may be arguable. In sentence final position, however, it does seem reasonable enough to believe that this is the case. It is apparent, too, that sentence (or clause) final intonation does have a distorting effect on lexical tones, and this effect, while manifested mainly in a change of register, can affect the shape of tones. The degree of change of shape is usually insufficient to cause ambiguity; however, this can and does happen. It seems obvious that the

degree of change would vary with the situation, and is to a certain extent dependent on the attitude of the speaker. A more extreme change would likely be associated with a stronger or more extreme emotion.

This chapter constitutes a review of experimental research done on the interaction between tone and intonation in tone languages, especially Chinese. It is a fairly new area of research and consequently the body of research done thus far is small. Most of the research has concentrated on the speech production aspect of this interaction, with a certain amount having been done on perception. General results have indicated that the shape of contour tones is not usually distorted in a significant manner, though the register of a tone may change. This is not the case in sentence final position, however, where intonation can and does have the effect of perturbing the shape of tones. It is possible that the perturbation be sufficiently large that tonal contrasts, such as between tones 2 and 3, may be neutralized.

II. Methodology

Evidence from speech production research has shown that tones can be distorted as a result of an interaction with intonation. Perceptual studies have indicated that this distortion can result in the neutralization of tonal contrasts. As a result, the present study was undertaken in an attempt to determine how much tones in sentence final position could be distorted and still maintain their identity. As shown in the previous chapter, it is in this environment that tones are most susceptible to perturbation.

The present chapter is a description of the methodology used for experimentation, of the response sheet used, and the make up of the subject pool.

A corpus of sentences was drawn up utilizing the sentence frame of Ho (1977): (in Mandarin, using Pinyin orthography) zhe ge zi nian ___. (This word is pronounced ___.) This sentence frame was used on the basis that, first, it can be spoken with a number of different intonations, and second, each word has a tone four, which allows for control of tone sandhi. For the sentence final position, the syllables 'bi' /pi/, 'ba' /pa/ and 'du' /tu/ were chosen. There were two main criteria for this choice: first, that each exists as four separate lexical items distinguished only by tone, giving a total of 12 test words, each of which is in common usage; and second, the three extremes of the vowel triangle are represented. Since common words using all four tones do not exist for all vowels in MSC, it was felt

that using the three extremities of the vowel system would be the best alternative. (See Table 1 for a list of test items.)

Six of the twelve sentences were recorded by a native speaker of the Peiking area dialect of Mandarin (MSC), using a Sennheiser MD 421N microphone and a TEAC A-7030 tape recorder monorally in a sound-treated room, at a speed of 15 ips. The fundamental frequency of the test items (sentence final syllables) was examined by means of oscillograms produced by a Frokjar-Jensen TransPitch-meter and recorded on a Mingograph 34. The oscillograms were segmented and measurements made at the beginning, middle, and end of the syllable nuclei of the test items. Mean values were calculated, first for the 4 tones across vowels, and then for each tone by vowel.

An analysis of variance was run on these values to determine the significance of variation between tones and replications. The ANOVA was done using three factors with replications. The three factors were:

Tone (T): the four tones of MSC, level, rising, falling-rising, and falling.

Vowel (V): the three vowels used were /i/, /a/, and /u/.

Segment (S): measurements were taken at the beginning, middle, and end of each token.

The results of the ANOVA are presented in Table Two, below, and are depicted graphically in Figs. 1a and 1b. Fig. 1a

This word is pronounced _____.
Zhe ge zi nian _____.

PINYIN	IPA	CHAR.	GLOSS
bī	/pī/	逼	to compel, to harass
bí	/pí/	鼻	nose
bǐ	/pǐ/	笔	pen
bì	/pì/	必	certainly, must
bā	/pā/	八	eight
bá	/pá/	拔	to pull
bǎ	/pǎ/	把	to control, to guard
bà	/pà/	爸	father
dū	/tū/	都	big city, capital
dú	/tú/	读	to read
dǔ	/tǔ/	赌	to gamble
dù	/tù/	渡	to cross

TABLE ONE: SENTENCE FRAME AND TEST ITEMS

SOURCE	ERROR TERM	SUM OF SQUARES	D.F.	MEAN SQUARES	F
TONE	R(TVS)	68639.83	3	22879.943	475.24*
VOWEL	R(TVS)	3928.06	2	1964.032	40.80*
SEGMENT	R(TVS)	9308.06	2	4654.032	96.67*
TV	R(TVS)	1904.38	6	317.397	6.59
TS	R(TVS)	113081.60	6	18846.934	391.47*
VS	R(TVS)	424.32	4	106.081	2.20
TVS	R(TVS)	2733.90	12	227.825	4.73*
R(TVS)		8665.83	180	48.144	

TABLE TWO: ANOVA RESULTS

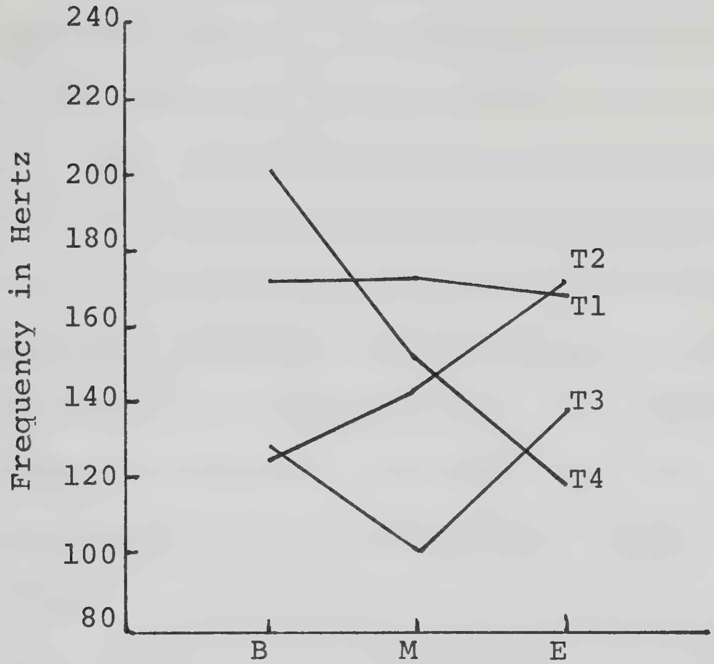

FIG. 1a:

Tone Position Averaged Across Vowels
((TS) Interaction)

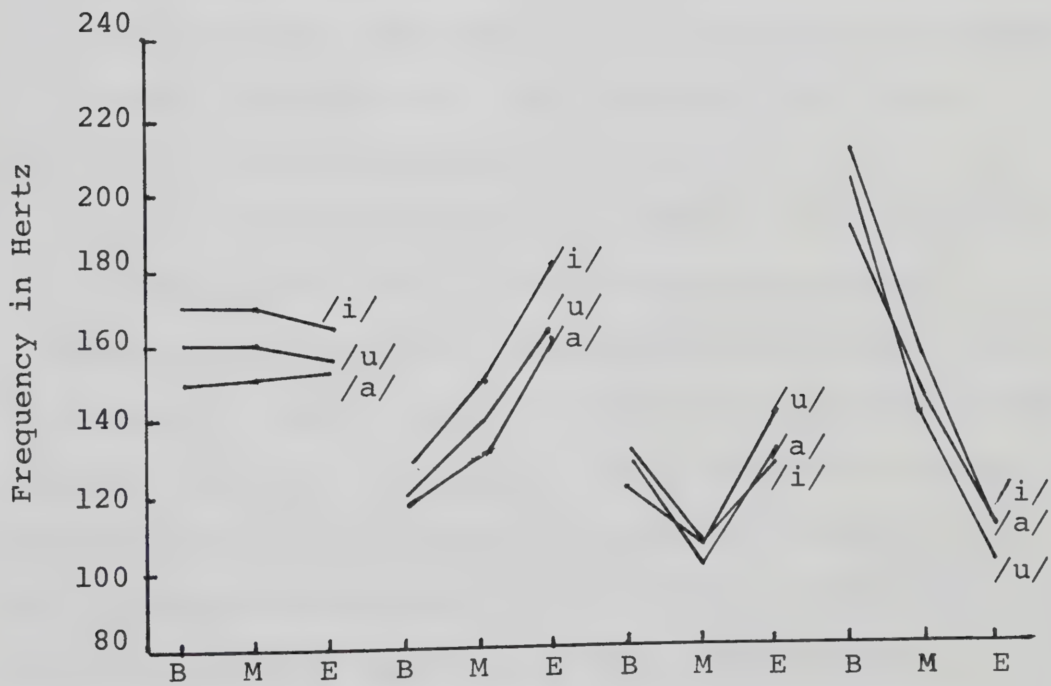

FIG. 1b: Tone Position by Vowel ((TVS) Interaction)

shows an averaging of frequencies for the four tones, for all replications of all vowels. The pattern here reflects the typical Mandarin tone patterns. Fig. 1b is a breakdown of this, depicting an averaging of replications for each tone by vowel and segment measurement. Significant differences are apparent between tones for each vowel. Introducing the vowel factor creates a slight shift in the patterning. The evidence that /i/ is spoken with the highest fundamental frequency, regardless of tone, then /u/, followed by /a/ is consistent with Lehiste (1970). This is true for all cases except for Tone 3, where /u/ has the highest fundamental frequency and /i/ has the lowest average beginning and end points. Tone 4 also shows slight variation, as /u/ has the lowest end point, rather than /a/.

Of the six replications, the fourth was chosen for use in the perceptual experiment as the fundamental frequency values for the sentence final syllables most closely reflected the mean values.

To assess the amount of perceptual difficulty that might come about, and to attempt to determine what 'confusions' among tones might occur, a gradient of 9 experimental conditions was established for each of the 4 tones of each of the 3 vowels. This involved manipulating the contour of the fundamental frequency of the final syllable of each sentence, a process described below. The desired intention was to construct a reasonable approximation of the effect of rising and falling

intonations, or any other prosodic feature characterized by changes in fundamental frequency, on lexical tones. The modified stimuli (as well as the natural, or unaltered tokens) were then presented to speakers of Mandarin in the form of a recognition task. (This is explained below.)

The various instruments used in this part of the experiment are described below:

1. Minicomputer; DEC PDP 12-A, word length 12 bits; 10 bit A/D and D/A converters; operating systems OS/8 and Alligator. The Alligator system is written in OS/8 PAL-12D and designed for use on a PDP-12 for the manipulation and presentation of stimuli used in psychoacoustic experimentation. (Stevenson and Stevens, 1978).
2. Tape recorder: TEAC A-7030 GSL. (Frequency response: 50-15000 Hz +/- 2dB; speed 15 ips; S/N ratio 58dB.)
3. Audio-frequency filter: Rockland 1524-01 (slope of frequency response:24dB per octave.)

The twelve sentences were bandpass-filtered to eliminate frequencies below 68Hz and above 6.8KHz, then sampled and digitized on the PDP-12, using the Alligator system. This set of sentences was then stored in an Alligator disc file. Using the Editor facility of the Alligator system, the final syllable of each of the 12 sentences was then truncated and stored in its own file. These syllables were then treated individually to create 9

experimental conditions for each, with the end-points varying by steps of 20Hz (see Fig. 2). Beginning points were held constant, and the total range of the gradient went from 80Hz to 240Hz. The sentence frame remained unchanged. Figure 2 presents sonograms of the 12 naturally produced syllables (3 vowels x 4 tones) and of the experimental conditions for ba (Tone 1). Also shown are graphs representing the experimental conditions for all tones and vowels.

Three separate programs were used in the construction of the stimuli; Extrac, Inton, and Patch (see Appendix). The program Extrac was designed to separate the syllable into individual glottal pulses, which were then stored in a file. The Inton program then used this file as a source file to call up the stored pulses (or vowel periods) individually and shorten or lengthen them, as desired, by a predetermined amount (see below). This amount was provided in another file, Pulse (see appendix for sample). The third program, Patch, then pieced together the adjusted pulses. The recreated syllables were longer or shorter than the original,and had to be corrected for duration. This was done through the Alligator editor, by deleting or adding individual pulses, as the situation demanded.

The technique for manipulating the stimuli is based on the inverse relationship between frequency and period duration of harmonic signals. The greater number of cycles the vibrating body (in this case, the vocal folds) makes in a specified amount of time (conventionally expressed per

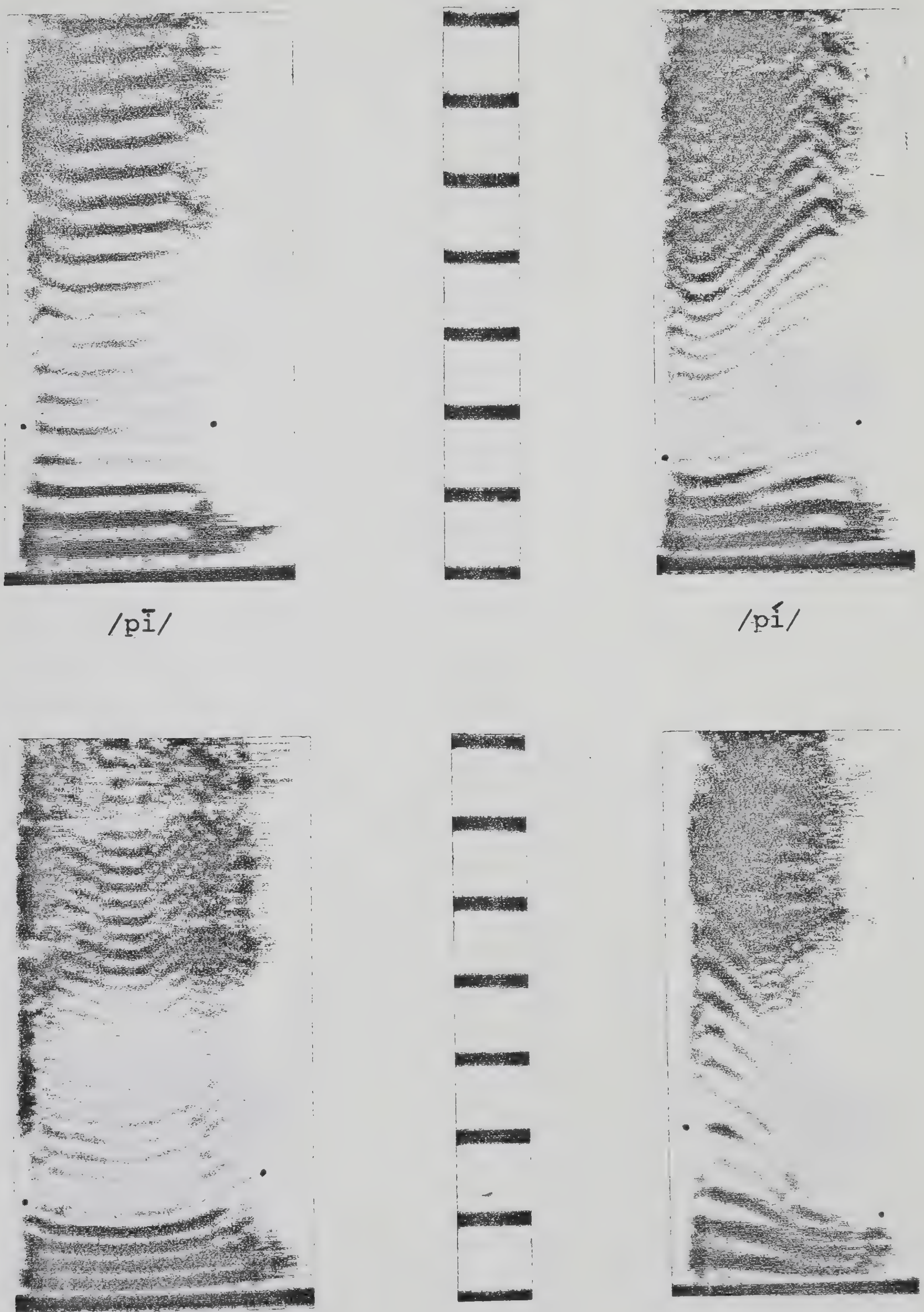

FIG. 2A: Sonograms of Naturally Produced Utterances

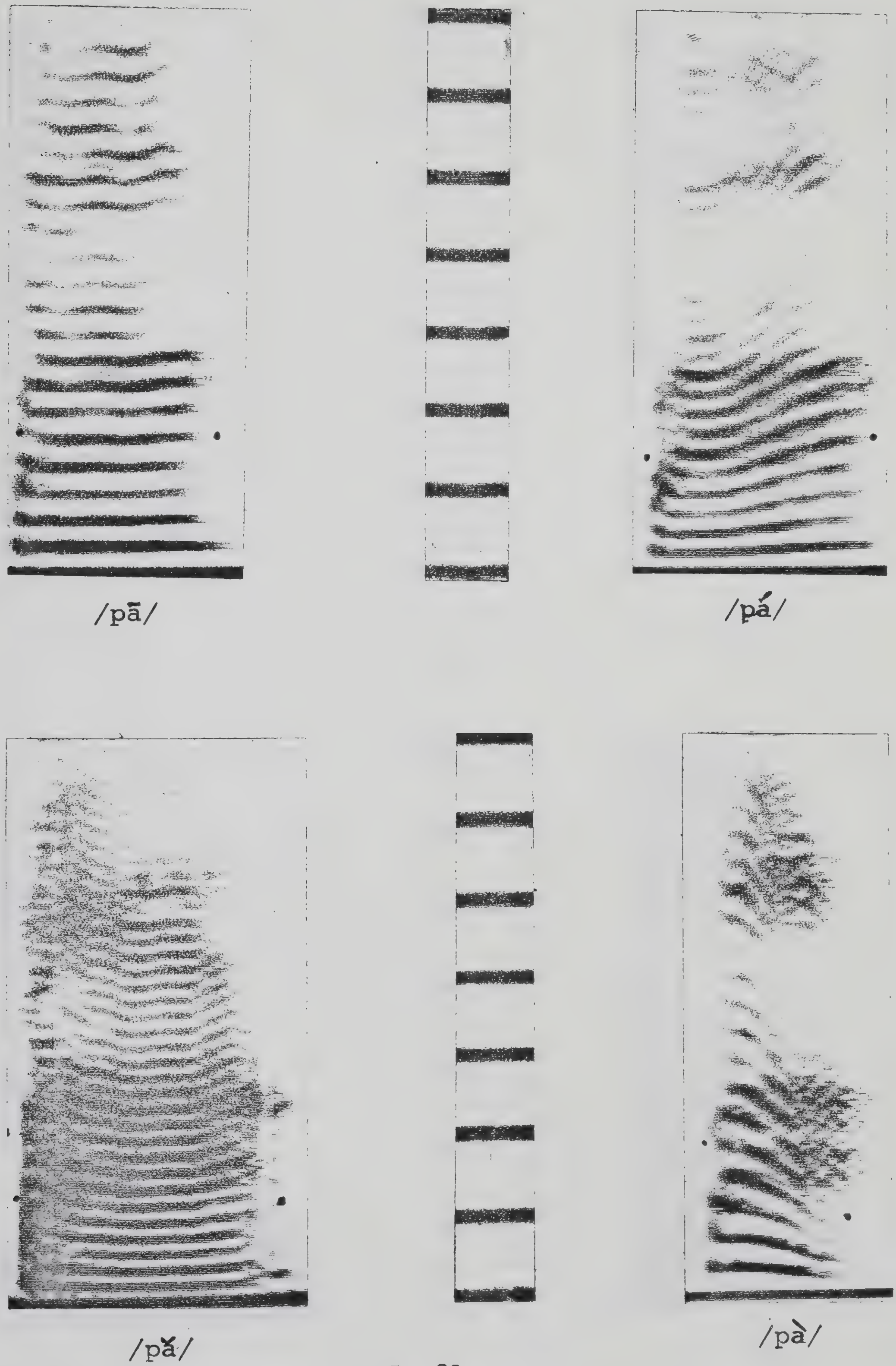

FIG. 2A

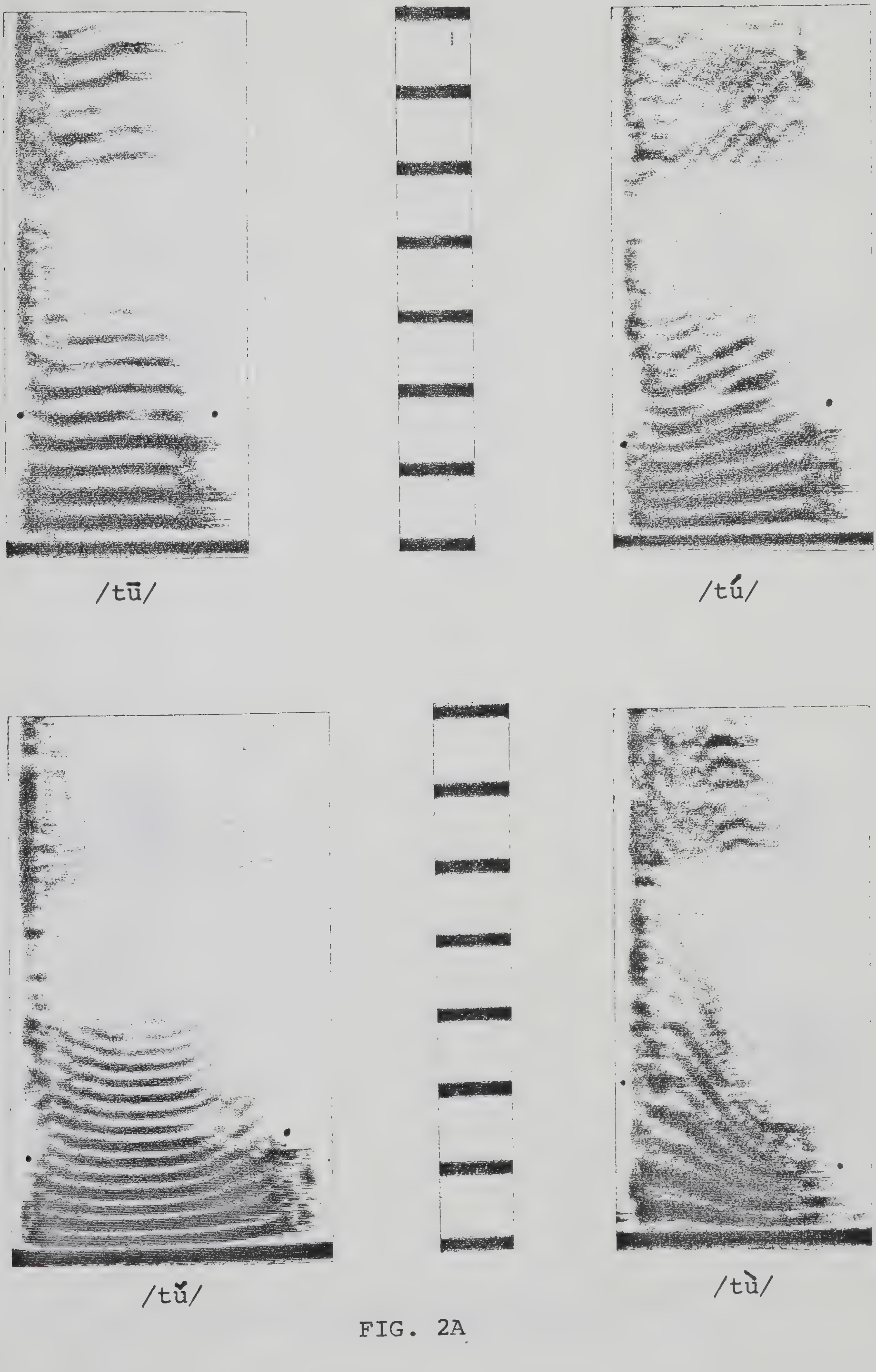

FIG. 2A

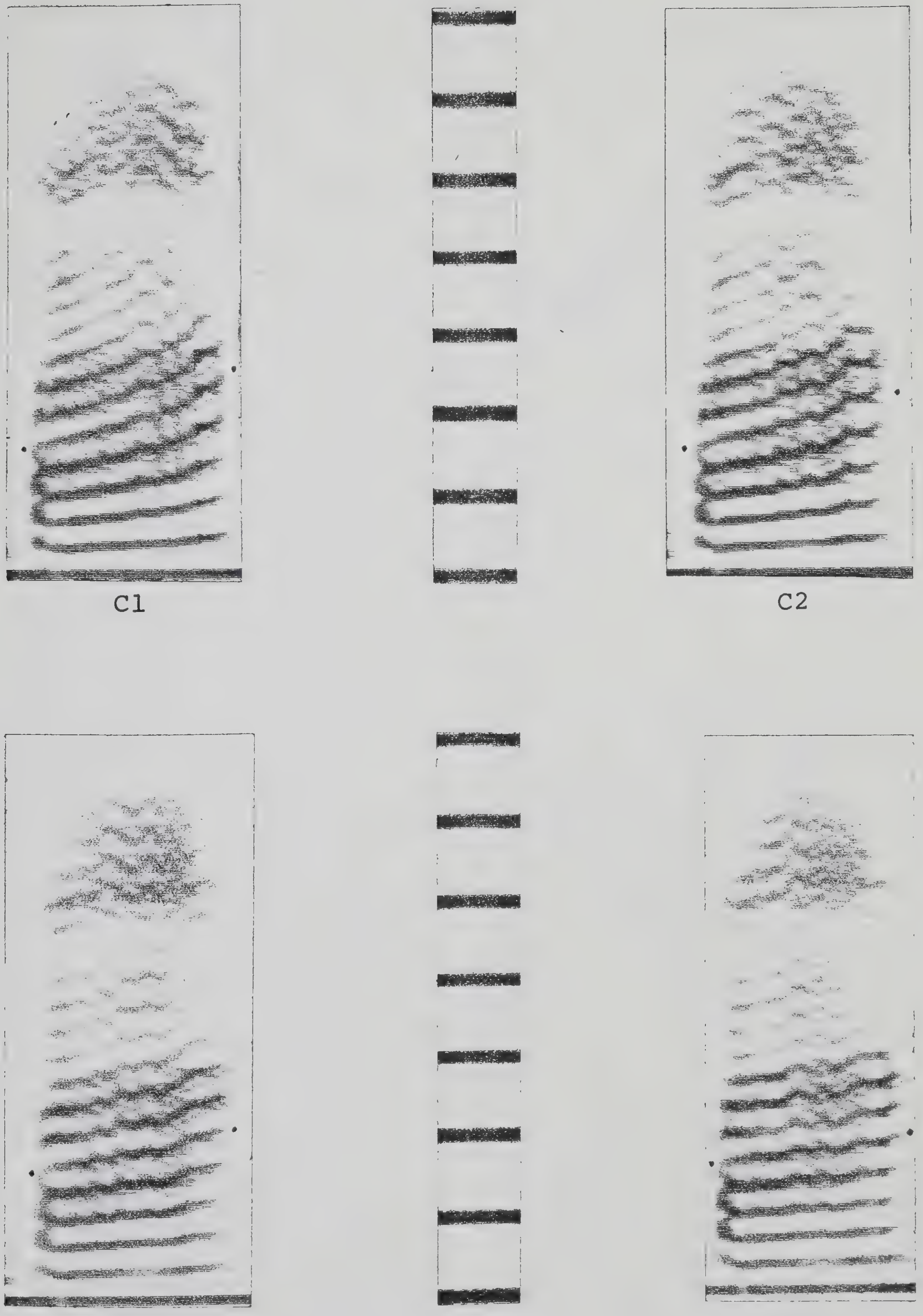

FIG. 2B: Sonograms of Experimental Conditions: /pā/

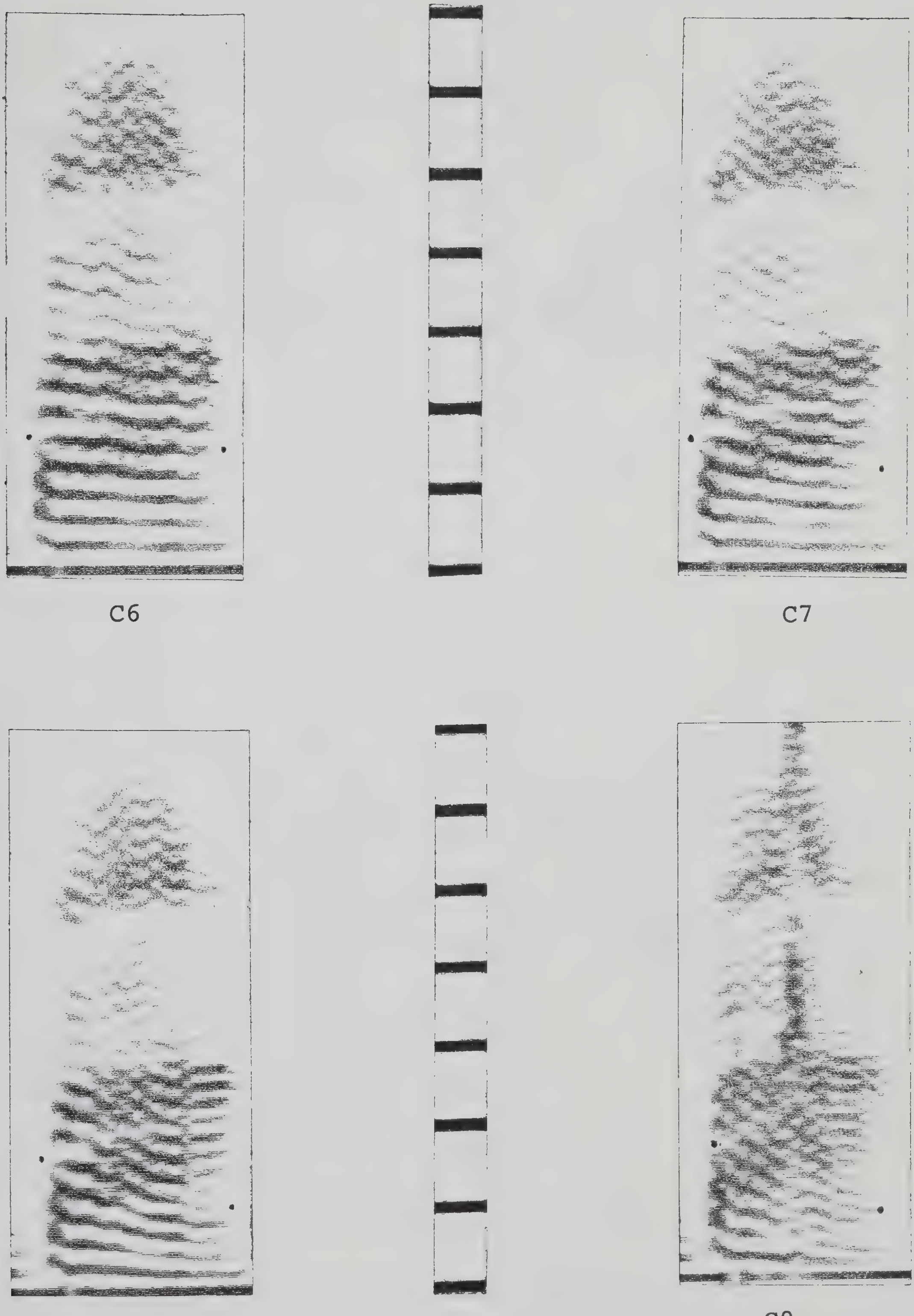

FIG. 2B

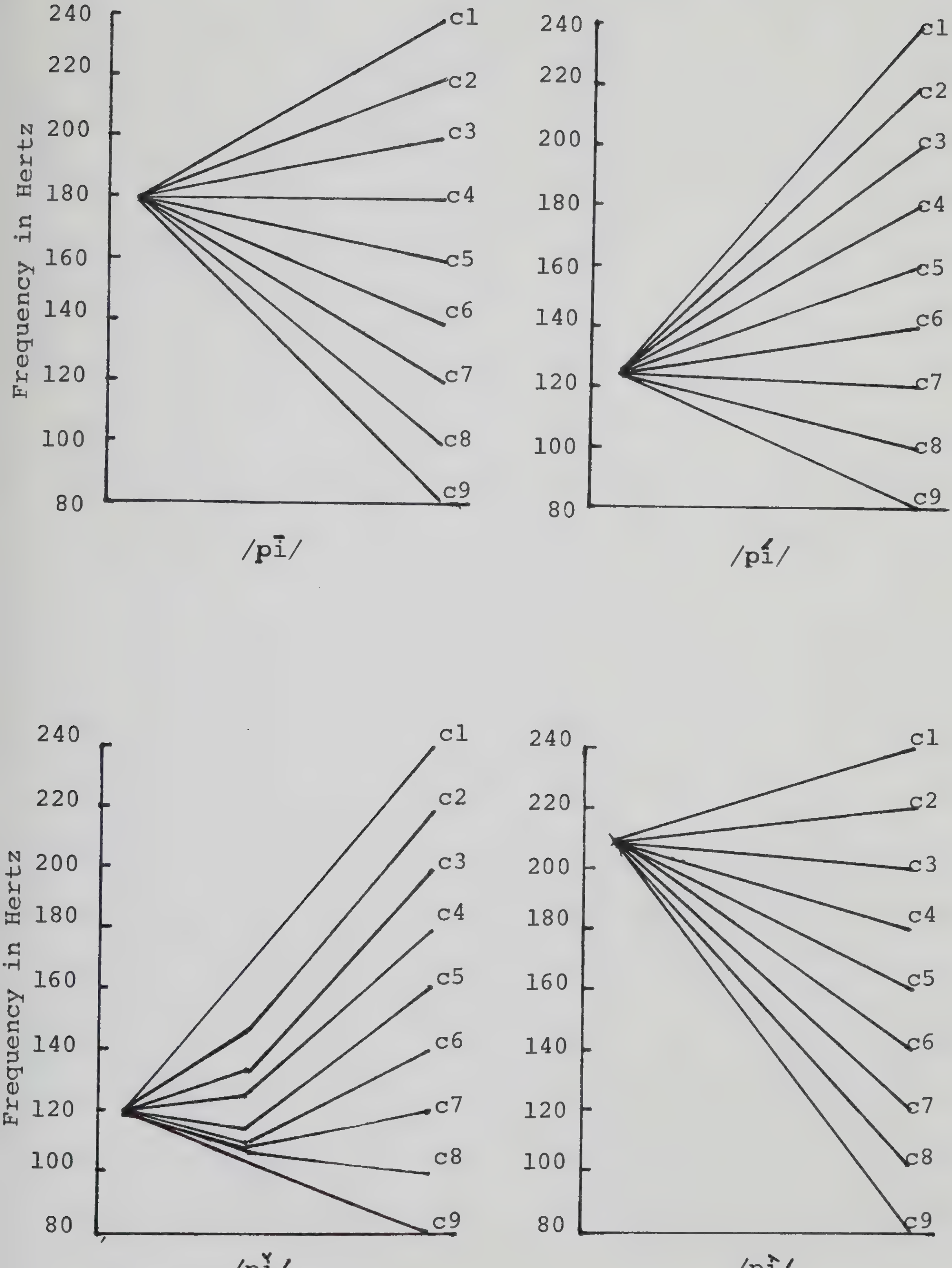

FIG 2C: Schematic Representations of

Experimental Conditions: /pi/

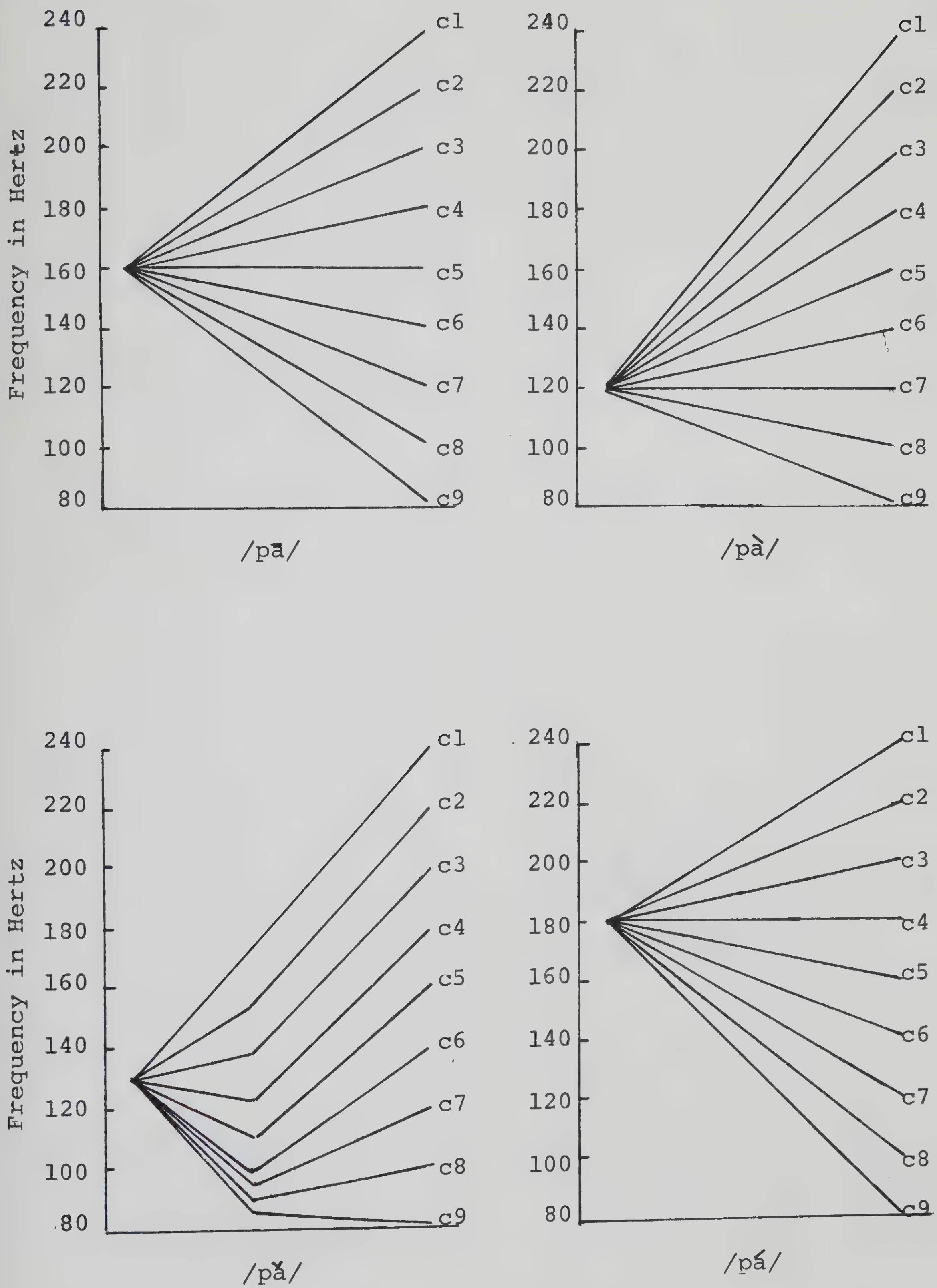
240
220
200
180
160
140
120
100
80
Frequency in Hertz
c1
c2
c3
c4
c5
c6
c7
c8
c9
/pā/
/pà/
/pǎ/
/pá/

FIG 2C: /pa/

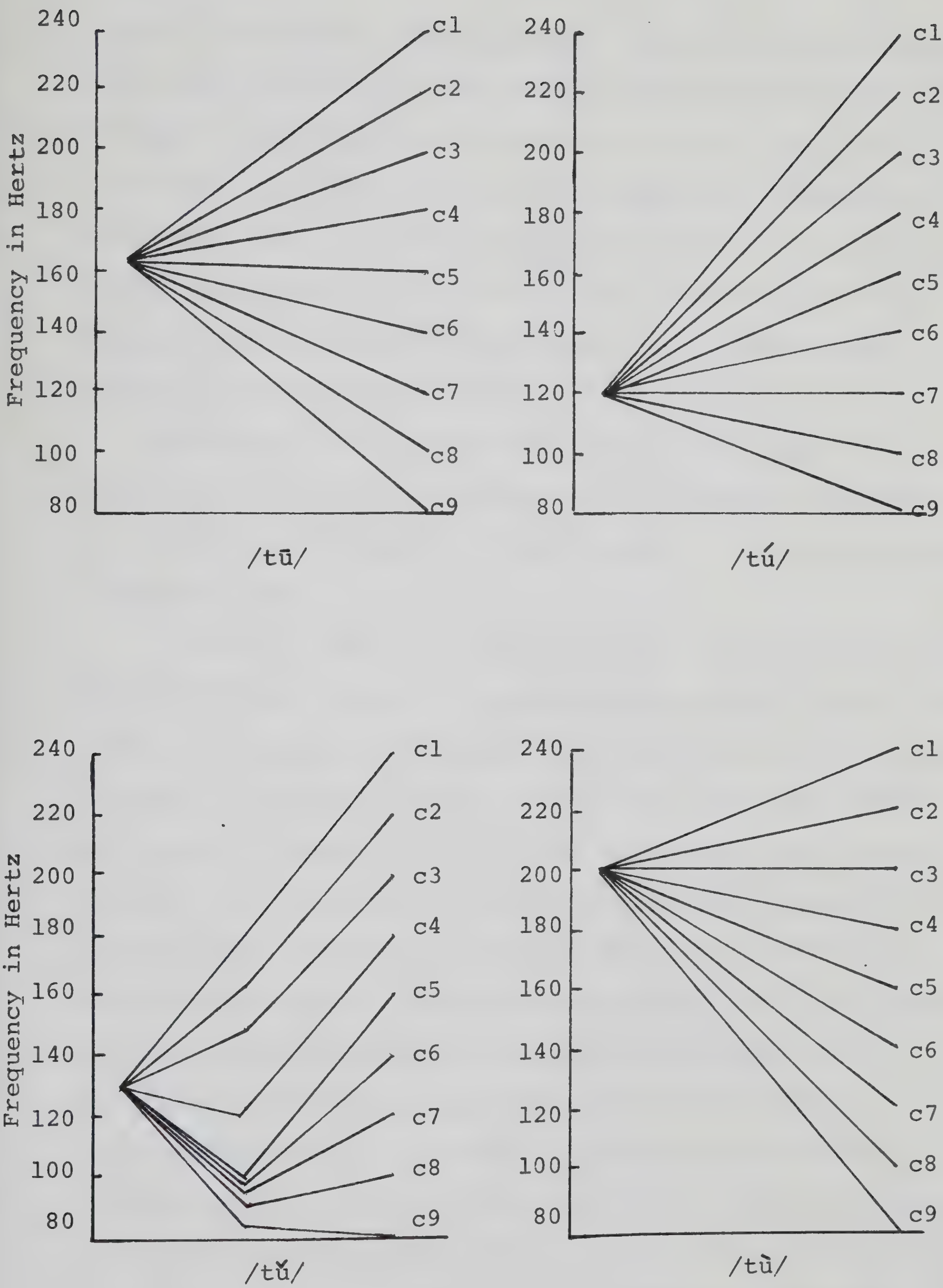
Frequency in Hertz
240
220
200
180
160
140
120
100
80
c1
c2
c3
c4
c5
c6
c7
c8
c9
/tū/
/tú/
/tǔ/
/tù/

FIG 2C: /tu/

second), the higher the frequency. (That is, each vibration has a shorter period.) For example, if the vocal folds take 10 msec to complete one vibration, or cycle, they will complete 100 cycles in one second. Hence the frequency of this sound (in this case, fundamental frequency, which refers to the basic pitch of a voice, determined by the vibration of the vocal folds) would be 100 cycles per second, or more commonly, 100Hz.

Therefore, it can be seen that by shortening the duration of individual pitch pulses, the number of pulses per second is increased, and an increase in the fundamental frequency results.

A sampling rate of 16KHz was used. At this rate, 16,000 data points are needed to encode one second of signal. For example, a signal of 100Hz having a period of 10 msec would consist of 160 data points; a signal of 140Hz, having a period of 7.1 msec ($T=1/F$), would consist of approximately 114 pts. By truncating 46 points from each period of a signal with a frequency of 100Hz, we can raise that frequency to 140Hz. A signal consisting of 40 pitch pulses can have a successively increasing amount of points subtracted from succeeding pulses, until the desired endpoint frequency is reached, resulting in a gradual increase in pitch. If the data points are removed at the end of the pitch pulse, where the signal is in decay, the vowel quality will remain unaffected. The wave was smoothed using a cosine squared window.

What has just been described is, in essence, the working of the program 'Inton', as set up to increase the pitch of a signal. To lower the pitch of a signal, points are added rather than subtracted. Points added had an amplitude of zero, essentially the addition of silence between each pulse. In extreme cases (e.g., lowering a steady frequency of 160Hz such that it is 160Hz at the beginning-point and 80Hz at the end-point) the technique again has its failings, as the increasingly large silence between pulses can have the effect of 'creaky voice', or perhaps to make the speaker appear to have a sore throat. This presented no real problems for this experiment.

Obviously these kinds of adjustments to the signal will have the effect of shortening or lengthening its duration. Since the duration of the different tones varies, and since the purpose of this experiment was to examine the effects of pitch variations on the perception of tones, it was desirable to control for duration. (Duration is a subsidiary cue for tone). This was done by adding or deleting entire pulses progressively throughout the signal, so as to retain a smooth sweep in pitch. In the case of adding pulses, this was done so that the same pulse was never (or as infrequently as possible) reiterated consecutively, to avoid the perceptual monotone or buzz that occurs with the reiteration of identical pulses.

In this fashion, the test items were created, then stored in disc files. Through the facility of the edit mode

the stimuli were rejoined to the original sentence frame. In each case this was done using the sentence frame corresponding to the original final syllables. For example, each of the conditions produced through manipulating 'ba' were joined to the sentence frame with which that syllable was originally spoken.

In total, there were 108 test items. (Three vowels x 4 tones x 9 conditions). Five different randomizations were produced through use of an Alligator program, giving a total of 540 tokens. These were passed through a desampling filter with a bandpass of 68Hz to 6,800Hz and then transferred to audio tape for presentation to the subjects.

The stimuli were presented to subjects through Telephonic TDH49 headphones. Subjects used an answer sheet (see Fig. 3) having Chinese characters to represent each of the 4 possible choices per stimulus, and were asked to circle the character corresponding to the stimulus presented. Instructions given to the subjects were presented in both English and Chinese, and questions were answered orally. The instructions were as follows:

> This is an experiment involving the tones of Mandarin. You will hear the sentence "This word is pronounced ___," with a different word at the end each time the sentence is played. Please circle the character on the answer sheet corresponding to the last word heard in each sentence. You will hear each sentence only once. Please begin with the left

1. 都 读 赌 渡

2. 逼 鼻 笔 必

3. 都 读 赌 渡

4. 逼 鼻 笔 必

5. 八 拔 把 爸

6. 都 读 赌 渡

7. 八 拔 把 爸

8. 逼 鼻 笔 必

9. 八 拔 把 爸

10. 逼 鼻 笔 必

11. 都 读 赌 渡

12. 八 拔 把 爸

13. 都 读 赌 渡

14. 都 读 赌 渡

15. 逼 鼻 笔 必

16. 逼 鼻 笔 必

17. 八 拔 把 爸

18. 逼 鼻 笔 必

19. 逼 鼻 笔 必

20. 八 拔 把 爸

FIG. 3: Sample Answer Page

column (Nos. 1-10) and work down, then continue with the right column (Nos. 11-20). After each 20 test items there will be a longer pause between items, giving you sufficient time to turn the page. You have approximately 2 seconds between sentences; after each 20 sentences there is a 4 second wait. There is a total of 540 test items and the experiment runs for 27 minutes. Are there any questions?

The first 10 items were presented as an example, without marking the score sheet, and then redone, and a 5 minute rest period was given at the half-way point. Subjects were also asked to indicate their age, province of birth, native dialect and other dialects of Chinese spoken.

The subject group consisted of 28 native speakers of Chinese. A subset of these (6) were natives of Taiwan, the rest were all natives of the Peoples' Republic of China and were at the University of Alberta as exchange scholars. It was not known in advance how many of the subjects were actually native speakers of Mandarin, though all claimed proficiency in Mandarin. Most of them, it turned out, spoke a number of different dialects.

The 28 subjects responded to 540 stimuli, giving a total of 15,120 responses. Due to variations in the subject group, such as native dialect, second dialects, and province of birth, there is some variation in the response patterns. Consequently, in the following chapters, the results are

analysed and discussed accordingly. Most of the discussion is of the results of a sub-group of the subject pool which consisted of ten people who were native speakers of Mandarin or had a native-like proficiency in the language.

This chapter has been a discussion of the instrumentation and methodology used to construct the test stimuli and conduct an experimental investigation into the effects of pitch variations on sentence final lexical tones of Mandarin. The following chapter presents a statistical analysis of the results of this experiment.

III. Statistical Analysis

Since the experimental results showed a certain amount of variability, and since the subject group was comprised of people of various linguistic and geographical backgrounds, as well as covering a wide range of ages, it was decided to first analyse the data from the point of view of determining whether subjects having similar backgrounds were reacting to the stimuli in a similar fashion. The possibility exists that the differences in background would prompt the use of different strategies in different subjects. Consequently, determining the existence of strategy groups must be the first step in analysing data of this nature, then to look at how the objects being tested were handled.

The technique of hierarchical cluster analysis as described by Ward (1963), Johnson (1967) and modified by Baker (Baker and Derwing, 1982) appears to be the appropriate statistical tool for testing for different groups within a subject pool.

Four steps were carried out for computing the subject clusters. Each subject would receive a score for each of the nine experimental conditions under the four tones and three vowel categories. Thus, each subject would have 108 (9 x 4 x 3) scores.

(1) The 108 scores can be arranged in a vector form in exactly the same way for each subject.

(2) The next step consists of constructing a coincidence matrix for each subject. This is done by a process similar

to taking an outer product of the subject vector with itself. An example of a typical outer product would be as follows:

$$(a\ b\ c) \times (d\ e\ f) = \begin{pmatrix} ad & ae & af \\ bd & be & bf \\ cd & ce & cf \end{pmatrix}$$

Special attention has to be given to the type of 'multiplication' used to produce each element in a coincidence matrix. For example, given the product:

$$(a\ b) \times (a\ c) = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$$

the element 'a' in the matrix is the result of a x a, and the zeros are the result of a x b, a x c, and b x c. This is to say that the operation between an element and itself just equals that element, and the operation between an element and any other element is zero.

(3) The third step is to construct a distance matrix between all pairs of subjects of subjects from comparing the pairs of coincidence matrices. The distance between subjects is calculated by counting the number of mismatches for the same positions in the matrices for the two subjects. Then the total number of mismatches is divided by the total number of elements in the matrix. This standardizes the distance between 0 and 1 and will be then of the matrix size.

(4) Finally, a distance matrix for all subjects is made and submitted to the cluster analysis technique with the use of Ward's method. (See also Wishart, 1978). The purpose of this approach is to evaluate the subjects' judgements of tones while not introducing any external criteria that would bias our interpretation of their abilities.

This method operates on the principle that at any given point in the analysis, the loss of information resulting from the grouping of individual subjects can be determined by the "total sum of squared deviations of every point from the mean of the cluster to which it belongs." (Everitt, 1974). All pairwise combinations of clusters are considered and the two clusters whose fusion results in the least increase in the error sum of squares (E S S) are joined. At the primary stage each individual is regarded as a separate group, i.e., the E S S is zero; the process is continued until all individuals are joined. The results can be described graphically by means of a dendrogram (see, for example, Fig. 4).

There is some difficulty with this technique in determining the significance of an apparent cluster (i.e., whether or not the cluster actually does exist), as there are no definite statistical criteria for this task. However, extensive randomization tests by Baker on the data used in that study (Baker and Derwing, 1982) have indicated that values of twice the mean distance for the total matrix constitute a greater amount of heterogeneity, whereas values

of less than half the mean show significant homogeneity.

The test for subject groupings were run separately on results from each of the 3 vowels; dendrograms indicating the clusterings are presented in Figs. 4a, 4b, and 4c.

Two distinct groups were found in regard to the treatment of /i/. The mean was established as 0.6. The linkage at 1.222 on the ordinate shows the joining of the two groups. In the case of the vowel /a/ the existance of more than one group was questionable, the highest link being a borderline case at 0.574. Consequently tests were done on these data, treating them first as one group, then as two groups (see below). For the vowel /u/ the uppermost link is at 0.376, therefore it seemed there was only one subject group within the pool in terms of the treatment of this vowel.(See Fig. 4c).

Following the establishment of groups within the subject pool it became possible to adequately examine the stimuli. This was done group by group. The expectation here is that the number of clusters found among the objects, for each vowel, should equal 4, since subjects were expected to respond in terms of 4 tones.

The following chapter will begin with a more in-depth look at the constitution of the subject pool, and then go on to examine how the stimuli were treated.

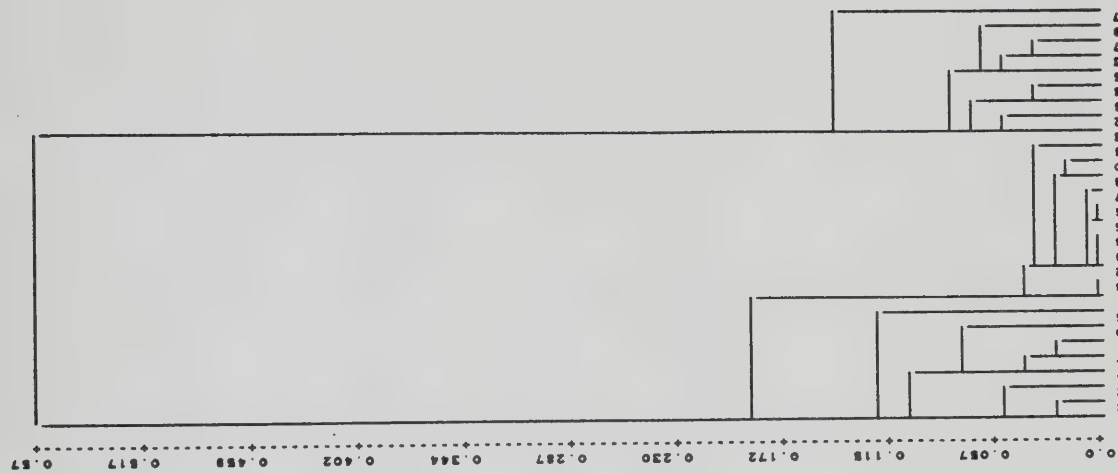

FIG. 4B

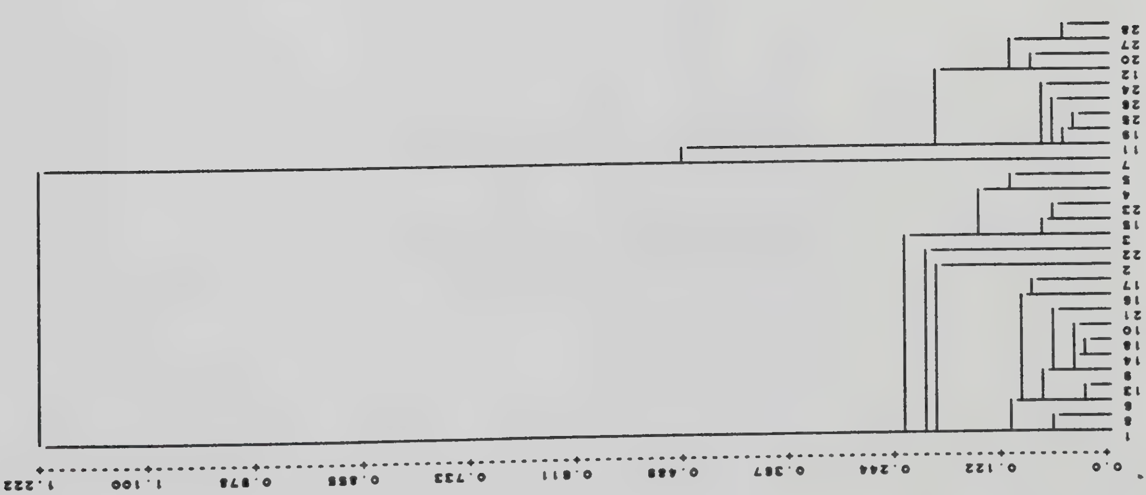

FIG. 4A

FIG. 4A: Subject Clusters /i/

FIG. 4B: Subject Clusters /a/

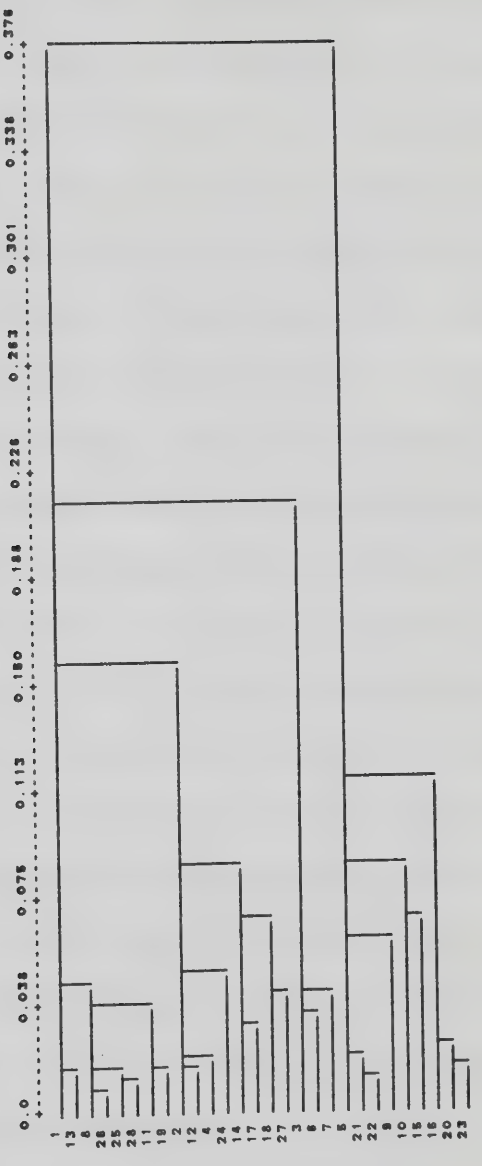

FIG. 4C Subject Cluster /u/

IV. Discussion: Subjects

In Chapter Three the lack of homogeneity among subjects was mentioned. In terms of this experiment, the most important aspect of this diversity is probably the variety of dialects spoken by the subjects. Of the 28 subjects, 9 were native speakers of Mandarin, 9 were native speakers of Wu dialects, and 3 were native Taiwanese speakers. Seven other dialects were also represented as native dialects for the remaining speakers. These are the dialects of the cities of Hofei, Chengdu, and Changsa, the provinces of Shensi, Shandong, and Gwangdong (the Hakka dialect), and the island of Hainan. Chengdu may actually be a Mandarin dialect. It is, however, considerably different from MSC, especially in its tonal system, and the Chengdu speaker declared MSC as his second dialect. Most of the subjects declared 1 or more second dialects or languages; however, of the 9 native Mandarin speakers, 8 of them answered negatively to the question concerning a second dialect. Of the remaining 20, those for whom Mandarin was not the first language, 14 indicated it to be their main second language. Two indicated Shanghainese as a second dialect, 1 Hunanese, and 3 did not indicate any second language.

Consideration of secondary dialects, or in the case of non-Mandarin speakers, first language, of the subjects is important in that possible interference from these dialects may have affected the responses of some subjects. Subjects were asked to respond to aural stimuli by circling a

character on the answer sheet.

The character writing system is common to all dialects of Chinese; there is, however a great deal of variation in tonal systems. The tone system of Mandarin (MSC) has four tones; level, rising, dipping, and falling, and is perhaps the simplest. Chengdu Mandarin also has 4 tones, but these are high-rising, low-falling, high-falling and low-falling-rising. Shanghainese has 5 tones, Cantonese has 6, 7 or 9 depending on dialect and analysis, and so on. Descriptions of the tonal systems of most of the secondary dialects found amongst subjects in this experiment are unavailable, so the amount of resulting interference is hard to quantify. Interference of this sort, however, is a possible contributing factor to the number of 'zero' responses (i.e. where less than 4 out of 5 instances of each stimulus were answered similarily - see preceding chapter).

Demographic considerations, and also age, are two other potentially critical factors. Mandarin constitutes the largest lingustic grouping in the Peoples' Republic of China, both in terms of numbers and speakers, and in geographical distribution. The dialect of the Peiking area is now treated as the national language (Modern Standard Chinese) and is becoming the predominent second language throughout the rest of China.

The geographical question is also important, in terms of the mobility of the population. Since the Communist revolution in 1949, a concerted attempt has been made to

break down the class structure of Chinese society. Partly this attempt has consisted in moving portions of the urban population - students, academics, bureaucrats, etc - to rural areas to live and work, and vice versa. This trend increased dramatically during the cultural revolution of the 1960's. The significance of this is that responses on the answer sheet to questions concerning province of birth and native dialect may not adequately reflect an individual's fluency in a given dialect. Native dialect is often understood as being the first learned; in such a highly mobile society this may not always be the one the individual is currently most proficient in, or has been most exposed to.

The question of age is also significant in this regard, in that those older subjects with Mandarin as a second language may be less proficient, having learned it at a later age, and therefore more susceptible to interference from a first dialect than those younger subjects who learned Mandarin as a second language at an early age.

Six of the subjects were not from the Peoples' Republic, but from Taiwan. Of these, 3 were native Mandarin speakers, having no second language; the other 3 are native Taiwanese speakers, but with a high degree of fluency in Mandarin. These three were included with the group of native speakers, since their responses were virtually identical, giving a sub-group of twelve listeners. Two of these were excluded from the following analysis, as their responses

showed considerably less consistency than did the others. The 10 remaining 'native' Mandarin speakers constitute what could be considered an 'ideal' group, in that they showed a high degree of consistency in their responses, both as a group and as individuals.

The treatment of the stimuli by each of the subject groups (as determined by the cluster analysis outlined in the preceding chapter) was also examined by means of a cluster analysis. For the group comprised primarily of native Mandarin speakers, identification curves based on the raw data were also drawn up.

The insights provided by the cluster analysis done on the stimuli matched expectations. For each vowel, the native group can be seen to have four distinct clusters (see Figs. 5a, c, e). Each of the clusters represents responses for one of the four tones. A strong correspondence can be seen between these and the identification curves presented below (Fig. 6).

Comparing the dendrograms for the native group with those of the rest of the subject pool (Figs. 5b, d) is of interest. While the native Mandarin speakers show a high degree of homogeneity in their responses, it is difficult to distinguish 4 corresponding clusters for the rest of the subject pool. This, too, could have been demonstrated through identification curves. However, these curves would have been so confusing, due to the lack of homogeneity in response patterns, that they have been omitted here.

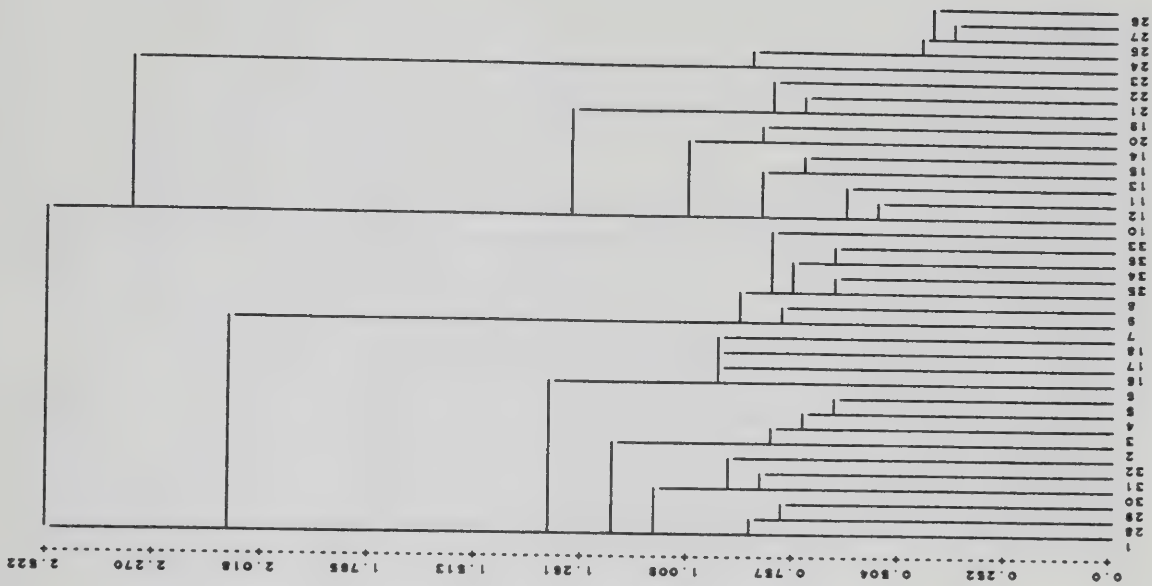

FIG. 5B

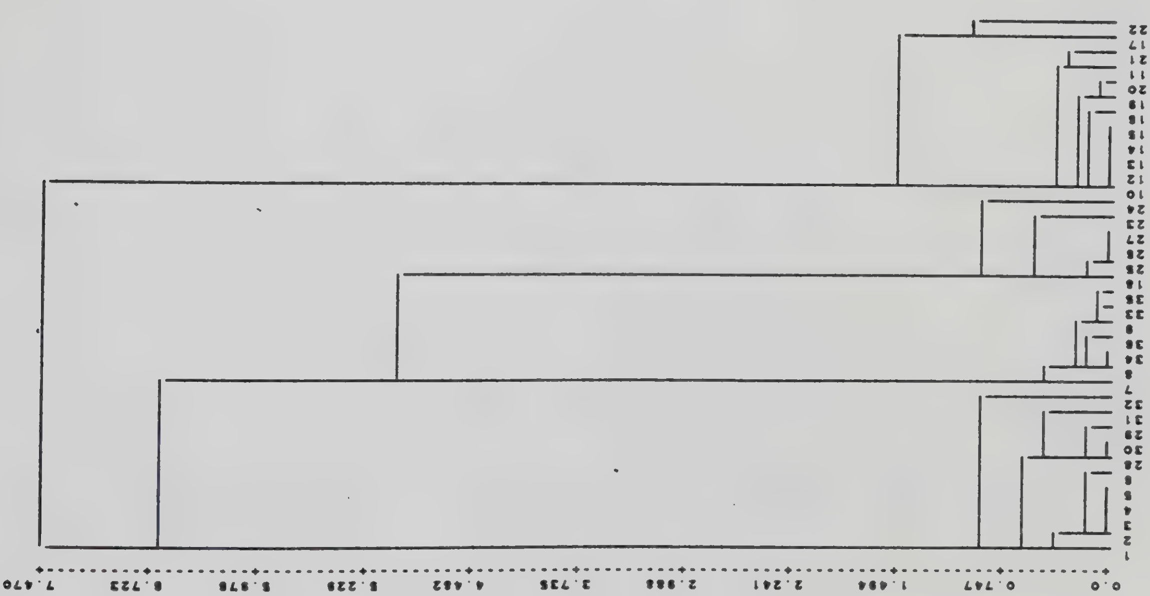

FIG. 5A

FIG. 5A: Object Clusters /i/ (Native Group)

FIG. 5B: Object Clusters /i/ (Non-natives)

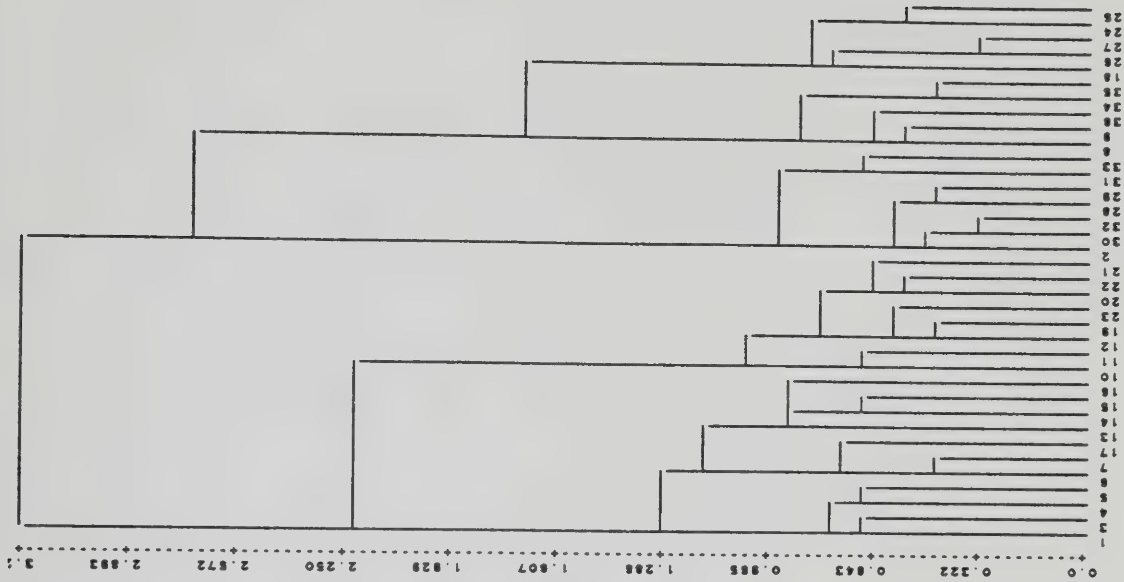

FIG. 5D

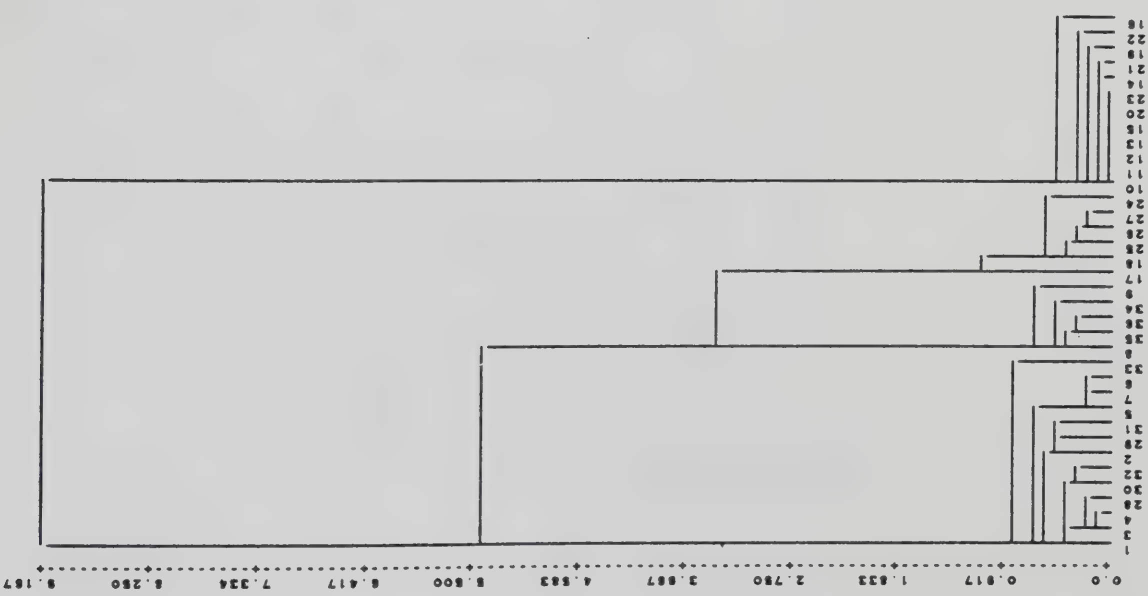

FIG. 5C

FIG. 5C: Object Clusters /a/ (Native Group)

FIG. 5D: Object Clusters /a/ (Non-natives)

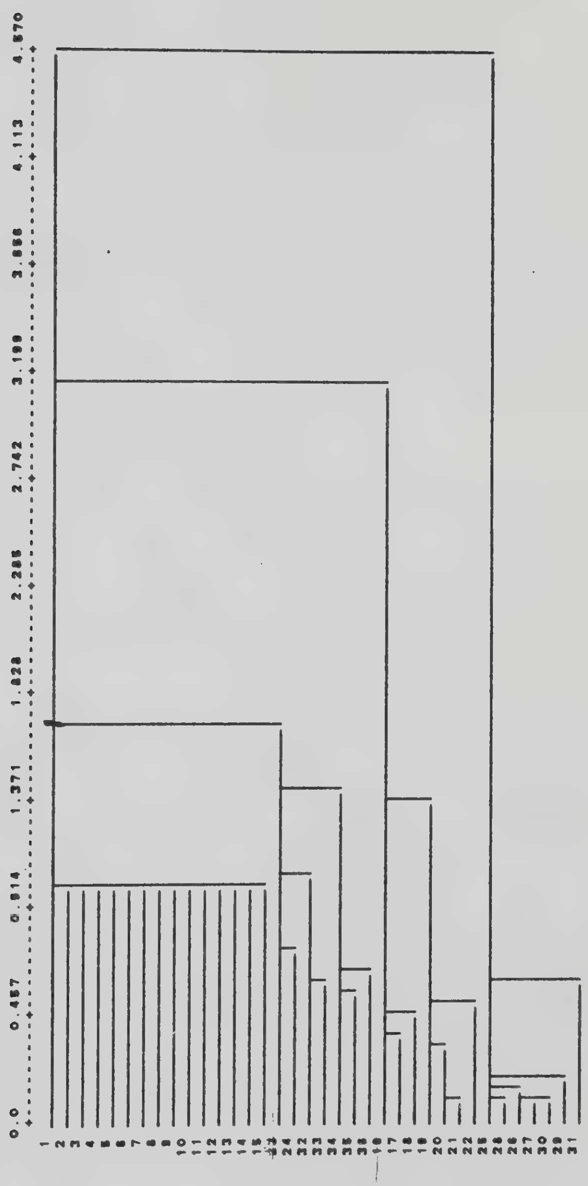

FIG. 5E: Object Clusters /u/

(All Subjects)

Identification curves describing the responses of the group of native Mandarin speakers are presented in Fig. 6. For the sake of clarity, identification curves for only two of the 4 possible choices are represented in each case. The solid line curve represents responses 'correctly' identifying the tone heard; the broken line represents the tone most often confused with, or substituted for, the original. Responses to the other two possiblities were for the most part negligible, and do not warrant consideration here unless otherwise mentioned.

The general trend of confusions, or substitutions, then, is as follows: Tone 1 is confused with Tone 4 when the end point is lowered by 40Hz or more from the original; raising Tone 1 (i.e., as in conditions 1-4) did not have the corresponding effect, i.e., Tone 1 could not be forced to be recognized as Tone 2. This statement is generally true; however, in the case of /tu/ there were more instances of confusion between Tone 1 and Tone 2. On condition 1 (the steepest rise) for /tu/ 36 respondants indicated Tone 1 and 14 indicated Tone 2. In the case of condition 1 for /pi/ and /pa/ there were 4 responses out of a possible 50 for Tone 2 and 45, in each case, for Tone 1. The pattern of these responses, then, seems to indicate that Tone 1, perceptually speaking, is not a level tone, as the production data indicates, but rather, a high tone.

A similar phenomenon can be seen happening at the lower end of the scale, though to a less pronounced degree. On the

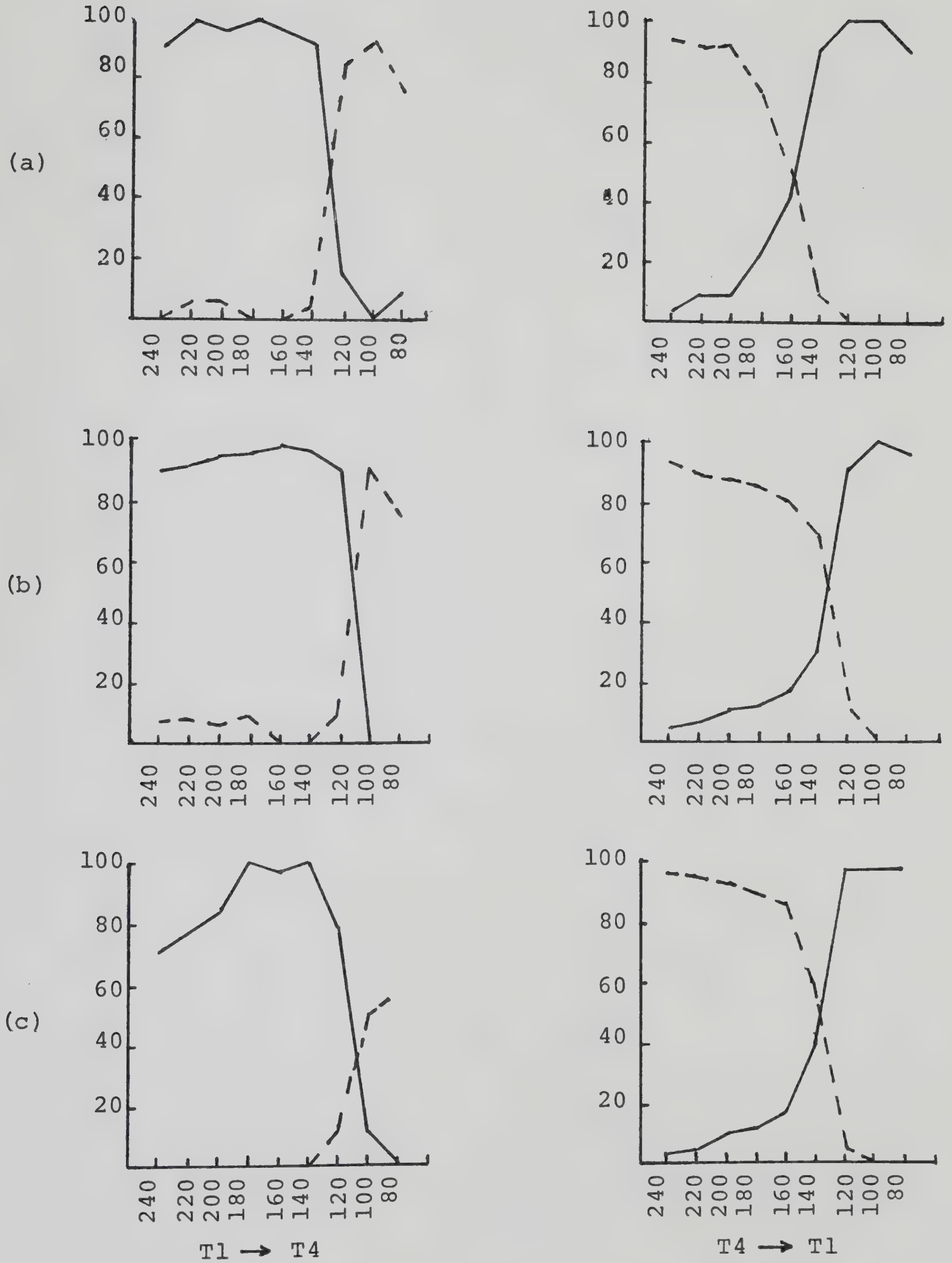

FIG. 6a, b,c: Identification Curves
ordinate: response percentages
abscissa: frequencies in Hertz

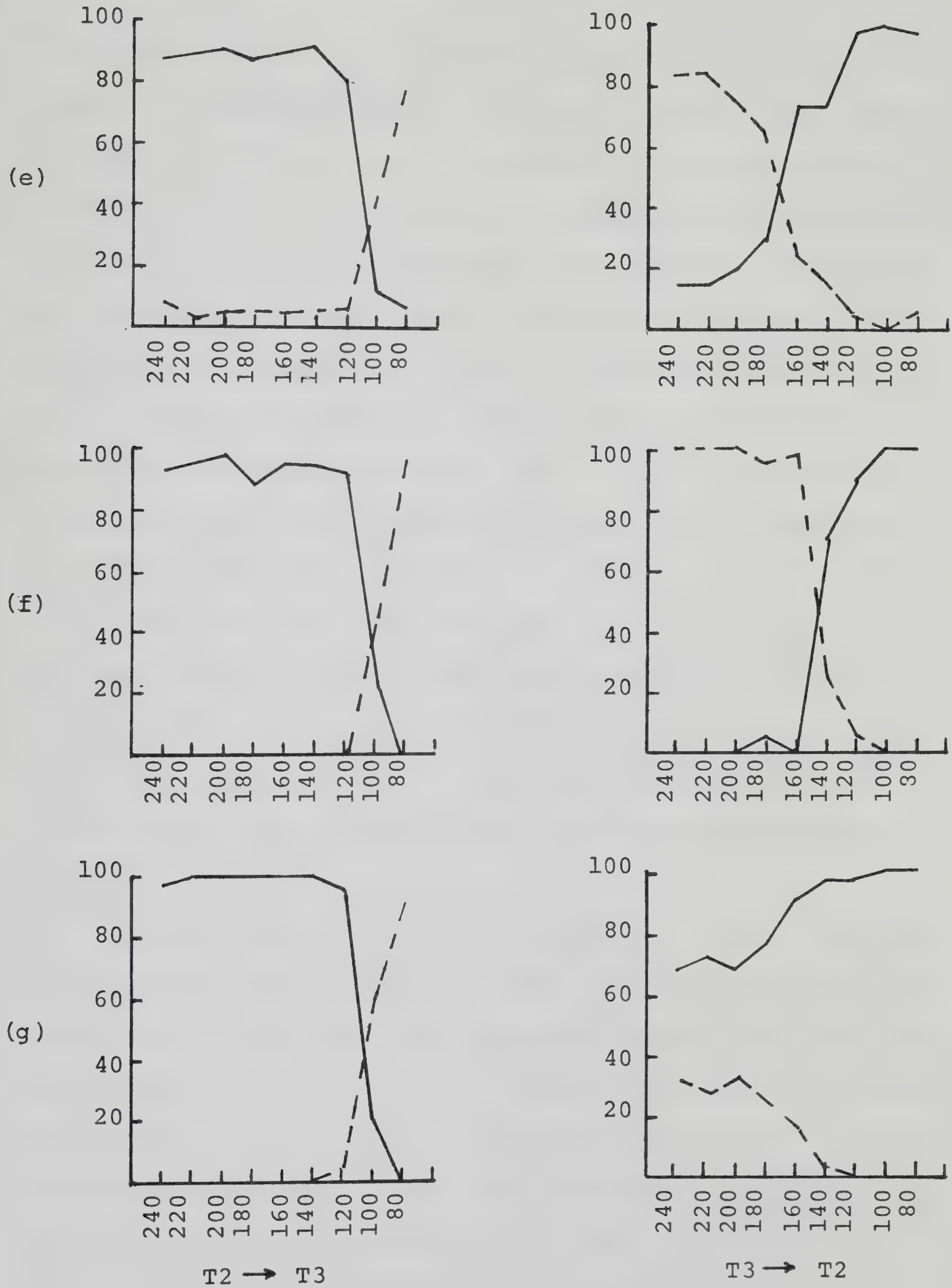

FIG. 6e, f, g: Identification Curves
ordinate: response percentages
abscissa: frequencies in Hertz

lower conditions (i.e., conditions 8 and 9) an increasing number of responses showing confusion of Tone 1 with Tone 3, as well as with Tone 4, is in evidence. This was strongest for /u/, with 19 of a possible 50 responses being identified as Tone 3. Twenty-nine reponses were indicated for Tone 4 and one each for Tones 1 and 2. The trend was less pronounced for /a/, with 12 Tone 3 responses and for /i/, with 8 Tone 3 responses. It could be argued then, from perceptual evidence, that MSC Tone 3 is a low tone rather than a dipping, or falling-rising tone as the production literature generally describes it.

Confusions resulting from perturbations of Tone 2 also correspond with evidence from the production literature. Lyovin (1978) reports on confusion of Tones 2 and 3 (see p. 20, above). Lowering Tone 2 to the extreme conditions (conditions 8 and 9) causes this tone to be recognized as Tone 3.

As demonstrated in the identification curves (Fig. 6e, f, g) recognition of Tone 2 is most consistent for /u/; the variations on the other two vowels for conditions 1 to 6 are very minor, it is significant, however, that for all three vowels the switch to Tone 3 occurs with condition 8. In this condition the natural Tone 2 had been manipulated such that the beginning points remained as natural (i.e., 125Hz for /i/, and 120Hz for /a/ and /u/), but with end points of 100Hz, and 80Hz for condition 9, thereby approximating a falling tone with no evidence of a 'dip' in the contour.

There were relatively few instances of Tone 4 being recognized; consequently it can be assumed that it is either the over-all lowness of the tone which precipitates recognition of Tone 3 or the starting point in itself.

A somewhat reversed situation obtains with the case of manipulated third tones. In these instances, but to varying degrees with each vowel, Tone 3 was recognized as a second tone. This trend was strongest for /a/ with 100% of responses indicating Tone 2 for the 3 upper conditions and only a slight drop for the next two conditions. A similar trend, though slightly weaker, is in evidence for /i/. For the vowel /u/, however, the incidence of Tone 2 recognition is much lower, never even reaching the 50% mark. It is also of interest to note that there were no cases of Tone 1 or Tone 4 being recognized for /u/, whereas there were a few for the other 2 vowels. Consequently there was a high degree of recognition of Tone 3 (i.e., minimal confusion) throughout all conditions. For instance, as mentioned above, for /a/, condition 1, there was 100% recognition of Tone 2; for /i/, under the same condition there was 82% recognition of Tone 2. For /u/ however, there was 32% recognition of Tone 2 and 68% recognition of Tone 3.

Again, as in the case of Tone 1 perturbations, it appears the tone can be pushed approximately 40Hz or more before confusion results. This was also true for Tone 2 perturbations, except for /u/, the perception of which did not change significantly until after a 60Hz drop.

The same situation is true of manipulations involving Tone 4, where confusions became predominant after 40Hz manipulation. In this case Tone 4 was increasingly recognized as Tone 1 after the end point had been raised by 40Hz. As indicated by the identification curves, there are only minor differences between the responses for the three vowels. The only noteworthy difference is again the greater consistency displayed for /u/.

The preceding is a description of how a particular sub-group of the subject pool reacted to the stimuli. This group was considered to be a group of native speakers of Mandarin because most of them actually were. Those who, in reality weren't, responded in a near-identical fashion. This group was also taken as the ideal, because of the high degree of consistency in their responses, as individuals as well as a group. In the preceding chapter we discussed the grouping of subjects by strategies, as determined by a hierarchical clustering analysis. It is now of interest to consider the 'native' group in terms of the subject groupings revealed by the cluster analysis. The essential questions are whether the native group is consistent with a grouping from the cluster analysis, and if so, does the cluster analysis agree that this group is indeed the most consistent in terms of responses. It should be re-emphasized that the cluster analysis, as done for each group and how they treated the stimulus objects, does not tell if subjects responded with the same answer, but rather if the degree of

coincidence of response is consistent within the group.

The treatment of /u/ will be considered first. In this case, it will be remembered, the cluster analysis revealed only one group, which obviously would include the native group. By examining the dendrogram for subject clusters for treatment of /i/ (Fig. 4a), it will be seen that, for the most part, the subjects constituting the native group (i.e., S's 1, 4, 11, 12, 19, 24, 25, 26, 27, 28) link together at relatively low points on the scale (about 0.015 to 0.02) indicating a high degree of homogeneity within their responses. The extent to which other subjects approach the linking points indicates the degree of homogeneity. In the case of the vowel /u/ (Fig. 4c), all subjects reacted in a similar manner, as indicated by the fact that the highest point of linkage is a 0.375. This is also reflected in the identification curves of the 'native' group for the tones of /u/, which show greater consistency than those of the other vowels. The fact that the different conditions did not always force a change in tonal recognition, as for /u/ in particular, is not reflected in the cluster analysis.

For treatment of tones on the vowel /i/, two subject groups were discerned by the cluster analysis, one consisting of 18 subjects, the other of 10 subjects. This second group is seen to be the more consistent of the two, and though not totally identical to the 10 'native' subjects, it is close as only 2 of the ten are different. Examining their treatment of the stimuli reveals that they

are consistent in their responses, as a group and as individuals. In terms of the coding system used on the raw data, which is described in the previous chapter (i.e., unless 4 out of 5 replications were answered the same a score of 0 was given; otherwise according to tone(1-4)), this group of subjects had no zero scores; the other larger group gave a majority of 0 scores for virtually all stimuli.

The difference between the two groups is reflected in the subject clusters for the two (Figs. 4a and 4c respectively). The fact that there are no distinct clusters for the first group shows there was no strong preference for any tone; the dendrogram for group 2 shows 4 distinct clusters, which proved to correspond to the four different tones, and are related to the identification curves for /u/, above.

A similar situation is evident for response to tones with /a/. There are again 2 groups discerned by the cluster analysis, one of 19 subjects and one of 9. The difference, however is that the native group (in its entirety) is contained in the larger group, rather than almost completely constituting the smaller group. The native speakers are again seen (Fig.**5c**) to be linked at lower points (generally speaking) than the other members of the same cluster, reflecting the greater consistency of those subjects.

As discussed in the previous chapter, the distinction of two clusters here was a border-line case. The cluster analysis concerning how each of these groups treated the

stimulus objects, however, supports this division. As is obvious from the dendrogram for group 1, this group treated the objects as four different groups. And again these corresponded to the four tones, and can be related to the ID curves shown above. The second group's dendrogram shows no clear cut distinction of clusters. An examination of the raw data shows a high proportion of '0-scores', indicating a higher degree of inconsistency in subject responses.

This chapter has been an examination of how the subject pool was divided into groups and how the stimuli were treated. The next chapter will look at these results in context of previous experimental research.

V. Discussion

The research reviewed in Chapter Two indicated that intonation can affect the shape and register of individual tones. It was also apparent that this occurs most strongly in sentence-final position. However, most researchers also pointed to a serious need for perceptual studies, to determine how serious the perturbation actually is. Specifically, this would be an attempt to discover the point at which a significant amount of confusion arises, or in other words, where the cross-over points of tonal categories would occur.

Furthermore, the perceptual work that has been done indicated the possibility of particular confusions arising. Most often the evidence seemed to show that confusion between Tone 2 and Tone 3 was likely: "...of the errors, most involved confusions between Tone 2 and Tone 3... both display rising glides, both start at about the same pitch level, and both have the same duration." (Gandour, 1978, p.45). Acquisition studies, both on infants (Li and Thompson, 1978) and on second language learners (Kiriloff, 1969) also point to the confusability of these two tones. It should be emphasized, however, that these studies all dealt with citation forms and consequently the results may not necessarily be extrapolated to the present situation.

Further support for the notion that Tones 2 and 3 may be confused comes from studies done on tone sandhi. One form of sandhi in Mandarin has Tone 3, when followed by another

Tone 3, approximating Tone 2. Psychological support may therefore exist for this particular confusion occurring in instances of a tone/intonation interaction.

There seemed to be less evidence to support the possibility of confusion between Tones 1 and 4, or 1 and 2; however these were considered possible in terms of a tone/intonation interaction, particularily in the sentence-final position.

The experimental stimuli were constructed in a manner that would allow for confusion to occur. That is, the conditions at both ends of the gradient (the range of experimental conditions) were considered to be sufficiently extreme that if perceptual confusions, or cross-overs of tone categories, actually could occur, the cross-over points would surely be included within this range of conditions.

As the results indicate, confusion of tonal categories do occur, though these didn't always match what might have been expected. As Figs. 6a, c, and e demonstrate, and as was discussed in the preceding chapter, Tone 1, when lowered by a sweep of approximately 40Hz, was largely recognized as Tone 4. On the other hand, when raised even by as much as 80Hz, Tone 1 was only very rarely recognized as Tone 2. This would indicate that the first tone should more properly be called (at least from the perceptual viewpoint) a high tone rather than a level tone. Futher research might confirm this suggestion. It also lends some credence to Rumjancev's (or Lyovin's) belief that the transitionsal portion between the

initial consonant(s) and the vowel contains sufficient information for tone recognition (see p.16, above).

In concurrence with observations on citation forms (above, p. 67) confusion does occur between tones 2 and 3. As the identification curves (p.59-60) indicate, this confusion works in both directions, Tone 2 being recognized as Tone 3, and Tone 3 being mistaken for Tone 2.

It is argued above that Tone 1 should perhaps be referred to as a high tone rather than a level tone; there is also evidence to suggest that Tone 3, generally referred to as a "falling-rising" or "dipping" tone could be called a low tone. The evidence is not as strong as in the preceding case, but is seen in Tone 2 going to Tone 3 in the extreme low conditions (especially for /u/), though Tone 4 still constitutes the primary object of confusion.

Regardless of what confusions actually do exist, it is evident that a fairly wide range of leeway exists before the possibility of confusion becomes problematic. Cross-overs do occur, and to a considerable extent, within the range used for experimentation. The degree of perturbation needed to create confusion, however, is quite large.

This wide range of flexibility, together with the use of particles to indicate functions often marked by intonation, helps lessen the possibility of ambiguity arising in ordinary conversation. Also, and of equal importance, these considerations leave sentence intonation free to perform other functions, as in other, non-tonal

languages, such as the attitudinal role of intonation.

Improvements on Experimentation

The greatest single flaw in the experiment of this thesis is the lack of homogeneity of the subject pool. All subjects had originally indicated a proficiency in Mandarin, however, the amount of regional variation found within this dialect was not anticipated. The mobility of much of the Chinese population, largely a result of the Cultural Revolution, was also not known beforehand. In North America, an alternative would be to work with Cantonese speakers, as these constitute the great majority of overseas Chinese. This however, would present other complications, especially in controlling for influences from English, as most of these people are bilingual.

The surest way, obviously, to ensure homogeneity of the subject pool in a situation such as this is to travel to China. In this way subjects could be pre-screened, but a pool of considerable size could still be maintained.

Another problem with the subject pool used in this experiment is the great variation in age. While in other circumstances this would not be reason for concern, in the case of subjects who are not native speakers of MSC it can be problematic. Attempts have been made since 1921 to make the Peiking dialect the national standard. It is only since the Communist revolution, however, that use of this dialect

has become widespread. Consequently, those older subjects who have learned MSC at a late age and possibly with no formal training will be more likely to suffer from interference from their native dialect. Younger subjects, having had exposure to MSC through school and the media, will be less influenced by this type of interference.

As far as the construction of the stimuli is concerned, there seems to be not much room for improvement at the present time. The biggest problem in constructing the stimuli is in accurately simulating the effects of intonation. There has not been enough research done on the shapes of different Mandarin intonations to know precisely the parameters involved. This is perhaps evident from the review chapter, where evidence was presented indicating that the sentence final intonation, rising or falling, occurs after the tone, whereas other researchers claim the intonation affects the entire final syllable. Still less is known about the manifestation of intonation in other parts of the sentence.

Future Experimentation

The present research has revealed a number of possible directions for future study. First would be a replication of this experiment, with an attempt to institute better controls over the subject pool. This, in all likelihood, would need to be carried out in China.

In the review of Rumjancev's work there was some discussion of his efforts to determine the relative importance of the question particle 'ma',vis-a-vis the rising question intonation. This is an important question, and more research should be conducted on it. The main problem involved was to first eliminate any trace of interrogative intonation from the particle; this was attempted by asking speakers to try to approximate a neutral intonation. The pitch manipulation technique used in the present experiment might prove more effective. Another problem encountered was that the particle 'ma' also has other meanings, for example, to indicate annoyance and this, too, is a confounding factor.

Other kinds of experimentation possible would include greater work on production and perceptual aspects of Chinese prosodics. Rumjancev has argued that apart from sentence final position, pitch may not be the primary cue for intonation. He maintains, rather, that duration and amplitude play a large role. There is a need, then, to establish the relative importance of these three parameters, from both the production and perceptual standpoints. Research could also be done to try to establish what part of the syllable carries the most relevant information in terms of tonal recognition. (See discussion on p. 16-17, above).

In short, there is room for a great deal of research in this particular area. Although some aspects of Chinese prosodics have been studied, for the most part this

research, while solid, is insufficient. Many important problems have yet to be examined.

VI. Summary

This thesis is a presentation of experimental research into the nature of the interaction of lexical tones and other prosodic aspects of Mandarin Chinese (Modern Standard Chinese). This area was chosen as a result of an interest in attempting to determine how problems resulting from use of the same acoustic cues for different functions would be overcome by users of the language.

Evidence in the experimental literature pointed to the lexical tones of sentence-final syllables as being the most susceptible to perturbation as a result of an interaction with intonation. Consequently it was decided to conduct an experiment designed to reveal how far the pitch of a tone (and hence the tonal shape) could vary before confusion as to tone identity would result.

A corpus of twelve naturally produced sentences was recorded. From these, the last (monosyllablic) word of each was extracted and used for signal manipulation for the creation of nine experimental conditions. These were constructed through the facility of a PDP-12 computer. The test items were then reconnected to the original sentence frames, and five replications were created.

The experiment was run on a group of twenty-eight exchange scholars from the Peoples' Republic of China. The subjects indicated how they recognized the sentence-final syllables by circling the corresponding Chinese character on the answer sheet.

A hierarchical clustering analysis was done to determine if different groups existed within the subject pool. Graphs are presented based on the results of a group of 10 subjects who were native speakers of Mandarin. The results obtained demonstrate that the kind of pitch variation used in the experiment (which approximates in a reasonable manner the influence of intonation on tones) can result in confusion of tonal identity. The most strongly occurring confusions were between Tones 1 and 4, where Tone 1, when lowered by a certain amount was generally recognized as Tone 4; and Tone 4, when raised by a similar degree, was identified as Tone 1. A similar situation obtained between Tones 2 and 3, i.e., a Tone 2, when lowered, was recognized as Tone 3, and Tone 3, when raised, was recognized as Tone 2.

The unexpected also occurred. It was assumed that Tone 1, when sufficiently given a rising shape, would be perceived as Tone 2. This in fact happened in only a very minor percentage of cases. It is argued, therefore, that the first tone should perhaps more properly be termed a high tone, rather than a level tone, which is usually the case.

To a lesser extent Tones 1 and 2 were recognized as a third tone in the lowest extreme. It is possible, then, that this tone is actually perceived as a low tone rather than as a falling-rising, or dipping, tone as it is described in the literature. Further experimentation could easily be done to test both these assertions.

Of the possible confusions, those caused by a rising intonation (Tone 3 becoming Tone 2, and Tone 4 becoming Tone 1) are seen as the most important. This is because, from a pragmatic standpoint, the actual occurrance of a rising, interrogative intonation makes Tones 3 and 4 susceptible to perturbation. The falling intonation associated with a normal, declarative sentence would not likely be severe enough to cause a change in tonal category.

The fact that Tone 1 could not be forced to be recognized as Tone 2, and Tone 3 as Tone 4, gives rise to the suggestion that the initial portion of the tone carries important cues for tonal recognition. Further experimentation could be done to test this notion. (See discussion of Rumjancev's work, in Chapter One.)

Whatever confusions resulted, it is apparent that lexical tones in MSC can withstand a large degree of perturbation before the danger of ambiguity arises. This leeway is crucial as it permits intonation to function in the same manner in Chinese as it does in non-tonal languages. Were the range of perturbability smaller, the use of pitch-oriented intonation for such linguistic functions as indicating interrogatives, and paralinguistic functions such as indicating attitudes and emotions, would be severely restricted. In this eventuality, the functions of intonation would conceivably be taken over by intensity and duration.

Certain problems were encountered with the experimentation, particularly concerning the subject pool.

It is believed, however, that these problems are unimportant in terms of the general trend of the results.

The nature of the relationship between the different aspects of prosody in Chinese is a new area for experimental research. The experimental studies reviewed in Chapter One, together with the experiment reported in this thesis have shed some light on this phenomenon. This body of research, however, rather than answering all the questions, has done much to open up a new area of research; there are a great many questions yet to be answered.

Bibliography

Abramson, A. S. The Vowels and tones of Standard Thai: acoustical measurements and experiments. International Journal of American Linguistics, 28/20, 1962.

_____. Lexical Tone and sentence prosody in Thai. Proceedings of the 9th International Congress of Phonetic Sciences 1979.

_____. The noncategorical perception of tone categories in Thai. In O. Creutzfeld, et al (eds.), Hearing Mechanisms and Speech. Berlin, Springer, 1979. p.127-134.

Baker, W. J. and Derwing, B. L. Response coincidence analysis as evidence for language acquisition strategies. Applied Psycholinguistics. 1982 vol. 3 no. 3

Baum, Richard, (ed.). China in Ferment: perspectives on the Cultural Revolution. Englewood Cliffs N.J. Prentice-Hall 1971.

Bolinger, D. Intonation. Harmondsworth, Middlesex, Penguin Books Ltd. 1972.

_____. Intonation across languages. In J. Greenberg (ed.). Universals of Human Language, Vol. 2, 1978, p.471-524.

Chang, C. T. Tones and intonation in the Chengtu dialect. Phonetica, 1958, 2, p. 59-85.

Chao, Y-R. A system of tone letters. Maitre Phonetique, 1930, 45, p. 24-27.

_____. Tone, intonation, sing song, chanting, recitative, tonal compostiton in Chinese. In For Roman Jakobson. The Hague, 1956.

_____. A Grammar of Spoken Chinese. Berkeley, University of California Press, 1968.

_____. Chinese tones and English stress. In L. R. Waugh and C. H. van Schooneveld (eds.). The Melody of Language. 1980, p. 41-44.

Creutzfeld, O., Scheich, H., and Scheimer, C. (Eds.). Hearing Mechanisms and Speech: proceedings of E. B. B. S. Workshop, Gottengen, April 1979. Berlin, Springer, 1979.

Crystal, D. Prosodic Systems and Intonation in English. London, Cambridge University Press, 1969.

Everitt, B. Cluster Analysis. New York, Halsted Press, 1980.

Forest, R. A. D. The Chinese Language. London, Faber and Faber, 1948.

Fromkin, V. I. (ed.), Tone: a linguistic survey. NYC, Academic Press, 1978.

Fry, D. B. Prosodic phenomena. In B. Malmberg, (ed.). Manual of Phonetics. Amsterdam, North Holland, 1968, p. 365-410.

Gandour, J. T. The perception of tone. In V. Fromkin (ed.). Tone: a linguistic survey. NYC, Academic Press, 1978, p. 41-76.

Gandour, J. T. and Harshman, R. Cross-language study of tone perception. In D. Sankoff, (ed.), Linguistic Variation: models and methods. NYC, Academic Press, 1978.

Greenberg, J., (ed.). Universals of Human Language, Vol. 2: Phonology. Berkeley, University of California Press, 1978.

Ho, A. T. Mandarin tones in relation to sentence intonation and grammatical structure. Journal of Chinese

Linguistics, 1976(a), 4, p. 1-13.

_____. The acoustic variation of Mandarin tones. Phonetica, 1976(b), 33, p. 353-367.

_____. Intonation variation in a Mandarin sentence for three expressions; interrogative, exclamatory and declarative. Phonetica, 1977, 34, p. 446-457.

Howie, J. M. Acoustical Studies of Mandarin Vowels and Tones. Cambridge, Cambridge University Press, 1976.

Jacobson, R., Fant, G., & Halle, M. Preliminaries to Speech Analysis. Cambridge Mass., MIT Press, 1952.

Johnson, S. C. Hierarchical clustering schemes. Psychometrika, 1967, 32, p. 241-254.

Karlgren, Bernard. Sound and Symbol in Chinese (revised ed.). Hong Kong, Hong Kong University Press, 1971.

Kiriloff, C. On the perception of tones in Mandarin. Phonetica, 1969, 20, 63-67.

Kratochvil, P. The Chinese Language Today: features of an emerging standard. London, Hutchinson University Library, 1968.

Lehiste, Ilse. Suprasegmentals. Cambridge, Mass. MIT Press, 1970.

Lehmann, W. P. Language & Linguistics in the People's Republic of China. Austin, University of Texas Press, 1975.

Li, C. and Thompson, S. The acquisition of tone. V. Fromkin (Ed.) Tone: a linguistic survey. New York, Academic Press, 1978, 271-284.

Lieberman, Phillip. Intonation, Perception and Language. Cambridge, Mass., MIT Press, 1967.

Lyovin, A. Review of Tone and Intonation in Modern Chinese by M. K. Rumjancev, Journal of Chinese Linguistics, 1978, 6, 120-168.

Maddieson, Ian. Universals of tone. In J. Greenberg (ed.), Universals of Human Language, vol. 2, 1978, p.335-366.

_____. An annotated bibliography of tone. UCLA Working Papers in Phonetics. 1974, 28.

Malmberg, B., (Ed.). Manual of Phonetics. Amsterdam, North Holland, 1968.

Newnham, R. About Chinese. Harmondsworth, Middlesex, Penguin Books, 1971.

McCawley, J. D. What is a tone language? V. Fromkin (Ed.), Tone: a linguistic survey. New York, Academic Press, 1978, 113-132.

Nooteboom, S. G., Brokx, J. P. L. and de Rooij J. Contributions of prosody to speech perception. IPO Annual Progress Report, 1976, 11, p. 34-54.

Ohala, J. J. Production of tone. V. Fromkin (Ed.), Tone: a linguistic survey. New York, Academic Press, 1978, 5-40.

Pike, K. L. Tone Languages. Ann Arbor, University of Michigan Press, 1948.

Rozsypal, A. J. Digital gating of speech signals. Language and Speech, 1976 19 57-74.

Rumjancev, M. K. Ton i Intonacija v. Sovremennon Kitajskom Jazyke. (Tone and Intonation in Modern Chinese.). Moscow, Izdatel'stvo Moskovskogo Universiteta, 1972. Reviewed by A. V. Lyovin, Journal of Chinese Linguistics, 1978, 6, p. 120-168.

Sankoff, D., (ed.). Linguistic Variation: models and methods. NYC, Academic Press, 1978.

Stevenson, D. C. and R. C. Stephens. Alligator Reference Manual, vols. I and II. Dept. of Linguistics, University of Alberta, Edmonton, 1978.

Vance, T. J. An experimental investigation of tone and intonation in Cantonese. Phonetica, 1976, 33, p. 368-392.

_____. Tonal Distinctions in Cantonese. Phonetica, 1977, 34, p. 93-107.

Wang, Wm. S-Y. Phonological features of tone. International Journal of American Linguistics, 1967, 33, p. 93-105.

_____. Language change. In The Annals of the New York Academy of Sciences, 1976, 280, p.61-72.

Ward, J. H. Journal of the American Statistical Association. 1963, vol. 58, 236.

Waugh, L. R., and C. H. van Schooneveld (eds.). The Melody of Language: intonation and prosody. Baltimore, University Park Press, 1980.

Wishart, D. Clustan User Manual (3rd ed.), Program Library Unit, Edinburgh University, 1978.

Woo, N. Prosody and Phonology. MIT thesis. Bloomington, Indiana University Linguistics Club, 1969.

Appendix

The following is a listing of the three programs used in constructing the stimuli used in the experiment. For details concerning their operation, the reader is referred to Chapter Two, and the Alligator Reference Manual (Stephens and Stevenson, 1978).

Extrac

```
GET D2:FILENAME
CLEAR ALL
SET NTH=4
SET VI=576
LQ TBA
DATA SEG*8
DATA Y
DATA T
DATA MEAS
DATA DF
ADJUST TBA ENTRY 0 &MEAS
LABEL 3
PR RESET SWITCH
READ *TTY &SEG &Y
IFERROR 43 LET &Y 16
ADD &MEAS &Y
ED
```

```
LABEL 1
WAIT 2SEC
IF #SS:1 EQ 0 GOTO 1
EX &SEG
RETURN
QUE &SEG
SAVE &SEG
MEAS DUR &Y &T
PR &Y &T
SUBTRACT &T 1
MULTIPLY &T 16
ADD &MEAS &T
PR &DF &MEAS
DISP TBA
ADJUST TBA ENTRY 0 &MEAS
GOTO 3
END
```

Inton

```
CLEAR ALL
SET NTH=2
DATA X*8
DATA T
DATA Y
DATA PTS
SOURCE D2:PULSE
```

```
GET D2:FILENAME
LABEL 1
READ *SOURCE &X &PTS
CL QI
LQ WIN
PR &X
LQ &X
WAIT 1SEC
MEAS DUR &Y &T PTS
PR &Y &T
ADD &T &PTS
LET &Y 628
SUBTRACT &Y &T
DISP WIN
WAIT 1SEC
ADJ WIN ENTRY 0 &Y
MOD MULT &X
MEAS DUR &Y &T PTS
PR &Y &T
ERASE *AF &X
SAVE SIG
RENS *AF SIG &X
DW &X
DW WIN
GOTO 1
END
```

Patch

```
GET D2:FILENAME
LQ P1
LQ P2
LQ P3
LQ P4
LQ P5
LQ P6
LQ P7
LQ P8
Q S1 P1 P2 P3 P4 P5 P6 P7 P8
GET D2:FILENAME2
SAVE S1
GET D2:FILENAME
CLEAR ALL
LQ P9
LQ P10
LQ P11
LQ P12
LQ P13
LQ P14
LQ P15
LQ P16
Q S2 P9 P10 P11 P12 P13 P14 P15 P16
GET D2:FILENAME2
SAVE S2
```

```
LQ S1
LQ S2
Q SIG S1 S2
SAVE SIG
PLAY
END
```

This program was extended or shortened, depending on the length of the signal being worked on. Most of the signals used had between 30 and 40 pitch pulses.

The following file was a source file, indicating how many points were to be truncated from each pitch pulse. It was used in conjunction with the program 'Inton'.

```
PULSE

P1 0
P2 0
P3 0
P4 0
P5 0
P6 2
P7 2
P8 2
P9 2
```

P10 2

P11 4

P12 4

P13 4

P15 4

P16 4

...etc. This file was adjusted, according to the number of pulses per signal and the amount of change desired in the pitch.


0 400 800 1200 1600 km
0 200 400 600 800 1000 mi
U.S.S.R.
U.S.S.R.
MONGOLIA
Heilungkiang
Harbin
Northern
Peiking
KOREA
Shanxi
Shungdong
Northwestern
Mandarin
Shenxi
Jiangsu
River
Shangai
TIBET
Szechwan
Chengdu
Hubei
Anhui
Wu
Zhejiang
Southwestern
Jiangxi
Hunan
Hun-anese
Gan
Min
Taipei
Tiawan
Gwandong
Hakka
Yue
Canton
Major Linguist Groupings
Dialects of Mandarin
National Frontiers
Provincial Frontiers